BOSTON COLLEGE
Department of Economics

EC 228: Econometric Methods (Fall 2017)
02: Stokes Hall 295S: T Th (9 – 10:15)
03: Stokes Hall 295S: T Th (10:30 – noon)

Christopher Maxwell
maxwellc@bc.edu
http://www.cmaxxsports.com

Maloney Hall, 337
Office Hrs: TTh 12:30 – 2
& by arrangement

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 project in which students will be building their own datasets and applying the various econometric methods developed in the course.

Peer tutors: Margaret Andersen & Kristian Singh @ Connors Family Learning Center

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 reference text:

- Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, Cengage Learning.

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), 5th (2012), or 6th (2015) 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.

- Angrist, Joshua D. and Jörn-Steffen Pischke, *Mastering 'Metrics: The Path from Cause to Effect*, 2014.
- Stock, James H. and Mark W. Watson, *Introduction to Econometrics*, 3rd ed., Addison-Wesley, 2014.

- Studenmund, A.H., *Using Econometrics: A Practical Guide*, 6th ed., Pearson, 2010.
- Ashley, Richard, *Fundamentals of Applied Econometrics*, Wiley, 2012.
- Kennedy, Peter, *A Guide to Econometrics*, 6th ed., Wiley-Blackwell, 2008.

Grading (Exams: 75%; Research project: 10%; Labs: 7.5%; Exercises: 7.5%):

- Three exams: Two mid-terms and one final exam
    - MT #1 (20%; mid-October… Oct 14th?): Ordinary Least Squares - SLR and MLR models
    - MT #2 (20%; Dec. 7th (the last class): Estimation and Inference + additional topics
    - Final (35%; 9:00 AM, Mon Dec 18th and/or 12:30 PM Tues. Dec 19th): covers full semester

  All exam grades are curved. In each case, I am hoping to have both sections of the course sit at one exam. Stay tuned.

  Exams are *open cheat-sheet* (two cheat sheets allowed for the final exam; one sheet for each mid-term).

- Final Exam: Optional. The Final Exam is optional. If you decide not to take the final, the weights on your mid-term exam grades are scaled up so that each mid-term exam counts for 37.5% of your course grade. You must commit to taking the final exam at the time you pick up that exam. To allow you to make a fully informed decision about whether or not to take the final exam, conditional course grades, assuming that you are not taking the final exam, will be posted to Canvas by the end of the day, Tuesday Dec 12th.

There are no make-up exams in this course. If you miss either mid-term exam, then you must take the final exam (exam weights will be adjusted proportionately).

- Research Project (10%): Replicate and improve upon an econometric analysis that has been published in an academic journal. The deliverable is a *PowerPoint* presentation, due Tues Nov. 21st, just prior to Thanksgiving Break. This will be a team assignment, with three students per team. I will assign teams. Grades will be given for both individual as well as collective performance. More details below.
- EC 228 Labs (7.5%): EC 228 course-wide labs, focused on using Stata in empirical/econometric analysis. Lab scores will be curved.
- Exercises (7.5%): There will be five Exercises over the course of the semester. 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. More details below.

Canvas: Boston College has in the past few years migrated away from BlackboardVista to a new course management system called ***Canvas***. That migration has not gone smoothly for my courses. Until Canvas offers *folders* (a radical new concept, I know), course materials will posted to the course website: http://www.cmaxxsports.com/ec228 . Note, however, that all exercise, exam, lab and research paper scores and grades will be posted to *Canvas*.

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.

Research Project (10% of total grade): This is an empirical research project, which will kick off with team assignments after the first mid-term exam. There are two phases to the project: you will first replicate an existing published piece of econometric analysis (of your choosing), [1] and then improve that analysis in some way (by adding more data, changing the specification of the model, changing the estimation technique, and so forth). Topics should showcase interesting econometric analysis, and need not be restricted to topics in Economics.

Your deliverable will be a *PowerPoint* presentation (or the equivalent). It should be concise and to the point; ***shorter is always better***. I will say more about the format of the deliverable when teams are assigned. Hardcopy presentations are due Tues Nov 21st, just prior to Thanksgiving Break.

- **Phase I : *Replication***

  Replicate both the summary statistics presented in the paper (to show that you have indeed replicated the construction of the dataset) as well as **at least one set** of regression results of interest. **Leave plenty of time for this phase. You'll find this far more challenging and time consuming than you could ever imagine.**

- **Phase II: *Improvement***

  Your turn! ... your improvement to the published analysis. This should be a lot of fun… but again, it will not go quickly or smoothly, so budget your time accordingly.

Your Phase I presentation should discuss your data sources and how your dataset was constructed; credit will reflect in part the level of difficulty.[2] In some cases you may be able to obtain data from the original authors, which obviously greatly simplifies the replication phase. You can do that if you want, but since building datasets is hard work, you won't get as much credit for your efforts as you would had you built the dataset yourself.

[1] Published here means published in an academic journal (so no unpublished senior theses, web blogs, or the like).
[2] If you want a sense of *degree of difficulty*, just ask.

In the Phase II portion of your presentation, you should brag about your outstanding improvements to the analysis, and discuss as well aspects of your analysis that did not pan out as expected.

I will assign the research project teams, which will likely have three members each. Students' grades will reflect both their individual performance as well as the quality of the final team product. *Peer evaluation* forms will be distributed at the end of the semester, so that team members can assess each other's performance.

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 or so one-hour sections of econometrics labs, taught by excellent BC Econ graduate student instructors. The lab sections will focus on the use of Stata in empirical analysis generally, and econometric analysis specifically. Lab scores, which will be curved, count towards 7.5% of your course grade.

Exercises: There will be five empirical exercises, counting towards 7.5% of your course grade. These will typically be team assignments (with 2-3 students per team) depending on how much work is required, and are typically graded on a 10 point scale. I will assign teams, which will differ from Exercise to Exercise. Teams should submit just one Answer Set. Final grades on Exercises are curved.

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. These will take some time to complete, so please do not leave them until the last moment.

Course Topics [Wooldridge 4th and 5th edition chapter numbers are in square brackets (there are just a few changes with the 6th edition)]

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

B: Introduction to Ordinary Least Squares (OLS)

2. Simple Linear Regression (SLR) Models [portions of Chapter 2.1-2.3]
    - *SLR analytics* (single explanatory variable): *in the beginning* (SLR.1: the data generation process); residuals and sum squared residuals (SSR); OLS as minimization of SSRs (FOCs and SOCs); sample statistics (sample means, variances, standard deviations,

covariances, and correlations); Sample Regression Function (SRF) and predictions; economic significance (elasticity)

- *SLR assessment*: Sum Squared Explained (SSE) and Sum Squared Total (SST); SST = SSE + SSR (w/ constant term in the model); goodness of fit (Coefficient of Determination, R^2), Mean Squared Error (MSE) and Root MSE (RMSE)

3. Multiple Linear Regression (MLR) Models [portions of Chapter 3.1-3.4]

   - *MLR analytics* (adding, and subtracting, explanatory variables): SLR analysis continued; interpreting coefficients I – *ceteris paribus* (partial effects and the SRF); the *collinearity* regression; multicollinearity; interpreting coefficients II – *what's new*; omitted variable bias (endogeneity)

   - *MLR assessment*: SLR assessment continued; adjusted R^2

Mid-Term Exam #1 about here

Sidetrip: Getting started with that research paper

4. Carrying out an empirical project [Chapter 19]

   - The scientific method: Testing hypotheses; collecting data; running regressions; doing it all again; and again; until… conclusions

C: Review of Probability and Statistics

5. Review of foundational basics: Probability and statistics [Appendices B & C] (some of this will have been covered in Part B. above)

   - 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 (Estimation and Inference): Populations; point estimators; sampling distributions; sample means, variances, covariances and correlations; unbiasedness; interval estimators; confidence intervals; hypothesis testing

     - Focus on estimation of the population mean, LUEs (Linear Unbiased Estimators), and BLUEs (Best Linear Unbiased Estimators)

D: Estimation and Inference in Regression Analysis

6. Estimation with SLR and MLR models [remainders of Chapters 2 and 3]

   - Estimation in SLR models: Gauss-Markov assumptions (SLR.1 – SLR.5); Population Regression Function (PRF); conditional means; means, variances, standard deviations

and standard errors of OLS estimators (intercepts and slopes); unbiasedness (OLS coefficients; MSE); LUEs; homoskedasticity; BLUE: The Gauss-Markov Theorem

- Estimation with MLR model (what changes? … not much!): … now MLR.1-MLR.5; Multicollinearity and Variance Inflation Factors (VIFs); n-k-1

7. Inference in SLR and MLR analysis [Chapter 4, portions]
    - Inference with SLR models: Add SLR.6 to the mix; normally distributed errors; variances, standard deviations and standard errors; t statistics; t-tests (Null hypotheses); p values; confidence intervals; hypothesis tests; economic v. statistical significance (elasticities v. p-values)
    - Inference with MLR models (what changes? … not much!): Now MLR.6; n-k-1; F-tests of linear restrictions; reported F stats and p values; Chow tests

E: SLR/MLR Analysis – Further issues etc.

8. Further Issues I [portions of Chapters 6-8]
    - Heteroskedasticity [Chapter 8]: Issues (OLS standard errors no longer correct; LUE but not BLUE); White-corrected standard errors (*robust* inference); working towards BLUE (weighted least squares… but where do those weights come from?)
    - Binary and categorical independent variables (Dummies) [Chapter 7]; fixed and interaction effects (intercept and slope dummies); average residuals (lazy economists); percentile dummies and more functional forms
    - Binary dependent variables [Chapter 7]: linear probability models
    - Functional forms (polynomials; logs; splines; etc) [Chapter 6]
9. Selected Topics [Handouts + portions of Chapter 15]
    - *Differences-in-Differences*: Deflategate; NBA Referee Own-Race Bias
    - *Regression Discontinuity Designs*: Highway Fatalities & Daylight Savings Time
    - Endogeneity and Instrumental Variables [Chapter 15]: The Oregon Health Insurance Experiment (Medicaid)
10. Further Issues II [portions of Chapters 5 & 6]
    - OLS asymptotics [Chapter 5]: Large sample properties; consistency (convergence in distribution)
    - Scaling effects and *beta* regressions [Chapter 6]
    - Forecasting and *prediction* intervals [Chapter 6]
11. Limited dependent variables [Chapter 17]: Maximum Likelihood Estimation (MLE); logit and probit models; censored and truncated regression models
12. More about specification and data problems [Chapter 9]
    - Misspecified models; proxy variables; measurement error in dependent and independent variables; missing data; outliers; non-random samples

Boston College
EC228: Econometric Methods

Calendar

I can't guarantee how exactly we will progress through the material. But based on previous semesters, here's a likely calendar (#s refer to weeks in the semester):

1. Syllabus; Introduction to the course; Econometrics examples

Intro to OLS: SLR and MLR models

2. *SLR Analytics*; Exercise #1 distributed (due wk #4)
3. *SLR Assessment*
4. Finish Intro to OLS/SLR; Exercise #2 distributed (due wk #7)
5. *MLR Analytics* (focus on differences from SLR analytics)
6. *MLR Assessment* (again, focus on differences viz. SLR assessment); Distribute practice questions for MT #1
7. Finish Intro to OLS/MLR; Mid Term #1 (Intro to OLS; Thurs Oct 14th ?)

Estimation and Inference

8. Kickoff Research Papers; Review of Probability and Statistics; Exercise #3 distributed (due wk #9)
9. Return MT #1s; Intro to MLR Inference (reading regression results); SLR and MLR models: Estimation and inference; Exercise #4 distributed (due wk #11)
10. SLR and MLR models: Estimation and inference cont'd; Research paper tips

Further topics/issues

11. Consistency; Heteroskedasticity; Dummy variables; Exercise #5 distributed (due wk #15)
12. More about dummy variables; Functional forms
13. *Thanksgiving Week*: Research papers dues on Tuesday
14. Selected topics; Distribute practice questions for MT #2
15. Semester wrap-up and review; Mid Term #2 (Thurs Dec 7th (last class); emphasis on material since MT #1)

Boston College
EC228: Econometric Methods

Stata @ 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.

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 … once CitrixReceiver has been installed on your computer. (If you are not connected to the BC network, you'll need to use VPN to access the apps server.) To learn how to access Stata through the apps server, go to http://www.bc.edu/offices/help/teaching/app_server.html .[3]

Alternatively, and to avoid traffic jams with Citrix and the apps server, you may want to purchase a six-month *Stata IC* license for $75 (sorry, but *small* Stata will not sufficient for Ec 228). For details, go to: http://www.stata.com/order/new/edu/gradplans/student-pricing/ .

We will devote some time to learning how to use Stata to build datasets and 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:

Encountering Stata questions/issues/features?… just *Google* it (always include "UCLA"). And here are a few sites that might be helpful (the pdfs are posted to the course website):

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

Examples and datasets (**bcuse** may be helpful here… I'll explain in class):

- 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/ecfindata.php
- http://www.ats.ucla.edu/stat/examples/

[3] And to learn about VPN, go to http://www.bc.edu/offices/help/getstarted/network/vpn.html .