BOSTON COLLEGE
Department of Economics

EC 228 – Econometric Methods (Fall 2019)
.02: T Th (10:30 – 11:45); .03: T Th (12:00 – 1:15);
O'Neill Library 257

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

Maloney Hall, 337
Office Hrs: TTh 2-4
& by arrangement

This is an introductory course in the use of econometric methods, with an emphasis on empirical applications. 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, categorical variables, functional forms 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 … 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: TBD @ 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*, 7th ed., Pearson, 2016.

- Bailey, Michael A., *Real Econometrics: The Right Tools to Answer Important Questions*, Oxford, 2016.
- Ashley, Richard, *Fundamentals of Applied Econometrics*, Wiley, 2012.

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

- Three exams: Two mid-terms and one optional final exam
  - MT #1: Thurs, Oct. 3rd OLS Analytics and Assessment
  - MT #2: Thurs, Nov. 21st OLS Estimation and Inference + Topics
  - Optional Final: date/time TBD … covers full semester

  All exam grades are curved.

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

  Optional Final Exam

  - If you decide to take the optional final exam, then each mid-term exam counts for 20% of your course grade, and the final exam counts for the remaining 35%.
  - If you decide not to take the optional final exam, 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 (optional) final exam, conditional course grades, which assume that you are not taking the final exam, will be posted to Canvas as quickly as possible.

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. More details below.
- EC 227 Labs (7.5%): Course-wide labs, focused on using Stata in empirical/econometric analysis. Lab scores will be curved.
- Exercises/Flips (7.5%): There will be five or so Exercises over the course of the semester. As well, there will be a few *Flips* (short for *Flip the Classroom*) and *q(uick)Flips* (very short flips). Course grades for Exercises/Flips 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/flip, 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.

Pass/Fail: It is perfectly fine, of course, to take the course Pass/Fail… but it is definitely not OK to do so and shirk on group projects/exercises. That is not fair to your teammates… and they will come to hate you! Accordingly: If you are taking the course Pass/Fail, please let me know at the start of the semester, and I will monitor goings-on and make adjustments if necessary.

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.

Topic selection: It's important to get off to an early start, as empirical research is always slow going!. To help you in that regard, I'll ask teams to email me a one paragraph description of their topic/paper of interest, by 6 PM on Weds. Oct. 23rd . I will compile those and we'll discuss them in class on Thurs. Oct. 24th .

- **Phase 0:** ***Topic Selection*** (briefly discussed in class, Thurs. Oct. 24th)

Replication and Improvement: Your end-of-semester deliverables will be two *PowerPoint* presentations (or the equivalent), which should be concise and to the point; ***shorter is always better***. I will say more about the format of the deliverables when teams are assigned.

- **Phase I:** ***Replication*** (due Thurs. Nov 14th ; three class-weeks after *Topic Selection*)

  Replicate both the summary statistics of interest 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*** (due Thurs Dec. 5th ; two class-weeks after *Replication*)

  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 your Phase II PowerPoint, 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. If possible, we'll have in-class research project presentations on the last day of classes.

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 products (the PowerPoint presentations as well as the in-class presentation).

Shirkers take notice: *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 (7.5% of your course grade): 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 will be curved.

Exercises and (q)Flips (7.5% of your course grade):

***Exercises*:** There will be five or so empirical exercises, which together with the *Flips* and *qFlips* (see following) count towards 7.5% of your course grade. These will typically be team assignments, usually with two students per team, 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/Flips 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.

***(q)Flips*:** We will have a few *Flips*/*qFlips* this semester. They can be *online* or *in-class* assignments, and are typically closely associated with material that we are covering in class.

- *Flips* will be graded on a four point scale. If they are in-class, they will be distributed in advance; students are expected to come to class fully prepared to *Flip*.
- *qFlips* will be very short Flips and typically an online outside-of-class assignment (mostly designed to give you some rudimentary practice with concepts and applications). They are graded on a two pt. scale; everyone should get perfect scores on these.

Boston College
EC228: Econometric Methods

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: *Way-Too-Fast* Review of Statistics

5. Estimation and Inference [Appendix C] (some of this will have been covered in B. above)
    - Our focus will be on estimation of the population mean, LUEs (Linear Unbiased Estimators), and BLUEs (Best Linear Unbiased Estimators)
    - Estimation: Populations; point estimators; sample statistics and sampling distributions (sample means, variances, standard deviations, covariances and correlations); unbiasedness; interval estimators
    - Inference: standard errors, t statistics, p-values; confidence levels; critical values; confidence intervals; hypothesis testing; significance levels

D: Estimation and Inference in Regression Analysis

6. Estimation with SLR and MLR models [remainders of Chapters 2 and 3; portions of Chapters 6, 8 & 9]
    - 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.
8. Related topics: Heteroskedasticity, F tests & measurement
    - 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?)
    - F-tests: Extension of t-tests to more complicated null hypotheses; testing linear restrictions with the F(q, n-k-1) distribution; relationship to changes in SSRs, R^2 and SSEs; connect the two approaches to Assessment (Goodness-of-Fit metrics and Inference); reported F stats (for the regression) and associated p values; Chow tests
    - Measurement – units and errors [Chapter 9]; scaling effects and *beta* regressions [Chapter 6]

E: SLR/MLR Analysis – Further issues etc.

9. Further Issues I [portions of Chapters 6-9]
    - Binary and categorical independent variables (Dummies) [Chapter 7]; fixed and interaction effects (intercept and slope dummies); average residuals (lazy economists)
    - Binary dependent variables [Chapter 7]: linear probability models
    - Functional forms (polynomials; logs; percentile dummies; splines; etc) [Chapter 6]
10. 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)
11. Further Issues II [portions of Chapters 5 & 6]
    - OLS asymptotics [Chapter 5]: Large sample properties; consistency (convergence in distribution)
    - Forecasting and *prediction* intervals [Chapter 6]
12. Limited dependent variables [Chapter 17]: Maximum Likelihood Estimation (MLE); logit and probit models; censored and truncated regression models
13. More about specification and data problems [Chapter 9]
    - Misspecified models; proxy variables; missing data; outliers; non-random samples

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://bcapps.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 $45 (sorry, but *small* Stata will not suffice for Ec 228). For details, go to: http://www.stata.com/order/new/edu/gradplans/student-pricing/ .

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

Boston College
EC228: Econometric Methods

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.

Stata Resources

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
- https://stats.idre.ucla.edu/stata/
- … and don't forget YouTube: https://www.youtube.com/results?search_query=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
- https://stats.idre.ucla.edu/other/dae/
- https://stats.idre.ucla.edu/other/annotatedoutput/

Also: Ben Lambert's *full course in econometrics* videos are terrific and come with a *British accent*!

- https://www.youtube.com/user/SpartacanUsuals/playlists