

**BOSTON COLLEGE**  
Department of Economics

**econ3370: Topics in Applied Econometrics (Fall 2024)**  
Fulton Hall 310: Weds (3:00 – 5:30ish)

Christopher Maxwell (maxwellc@bc.edu)  
Maloney Hall 337

TA: tba  
Office Hours: tbd

**Course Description:** This is an advanced econometrics/data-analytics course and a follow-on to econ2228. Selected topics include linear and nonlinear least squares estimation as well as a variety of maximum likelihood estimation topics including logit/probit, ordered logit/probit and poisson/negative binomial count regression models. A term-long research project is an important part of the course. The analytics in the course will be Stata- and Excel-based, but students, in doing their own work, should feel free to work in R, SAS, Python, Matlab or whatever their statistical language of choice might be.

Illustrative examples and datasets are drawn largely, but not exclusively, from the sports world. We could easily work with other data, but there's so much publicly available sports-related data available... Why not? ... and besides, it's fun! Non-sports-related topics/datasets that we'll likely consider include: pandemic propagation (e.g. covid), gun control, traffic fatalities, real estate valuation and S&P sovereign debt ratings.

**Course Philosophy:** I have two main goals in this course:

- The topics considered in this course are quite advanced, even for graduate students in economics. My goal is to give you a good sense (intuitive understanding) of those topics, relevant applications and associated tools of analysis, without getting bogged down in the often way-way-complicated math/stats. That goal is achievable!
- Data analytics is a skill, and you can never practice enough. So you'll be working with lots of datasets in a variety of formats. You will almost never be working with canned datasets. I want to give you something you can brag about in a job interview!

**Prerequisites:** Introductory Econometrics (econ2228) or an equivalent course such as Introduction to Data Science (csci2291). If you took econ2228, you are expected to have excelled in that course and received a metrics course grade of B+ or better, and preferably an A or A- ... and to know how to run basic econometric models using Stata and interpret regression results. If you took csci2291 and worked in Python, you should be able to pick up Stata quickly (as mentioned above, you can also continue to work in Python if you wish).

**BC's Minor in Data Science:** Some have asked if this course counts toward BC's newly established Minor in Data Analytics. My short answer: Not yet! ... mostly because I haven't asked. But there's no doubt that it should, and so I'll be working on that.

**Texts/Canvas:** There are no required texts for this course. I will post materials to Canvas as necessary.

**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 staff in the

## **econ3370F24: Topics in Applied Econometrics**

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, see BC's academic integrity policies webpage.<sup>1</sup>

**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.

**Class Structure:** This course has a somewhat unusual class structure. We will meet once a week for about 2.5 hours. The first half or so of class time (Half #1) will be devoted to standard pedagogy. The second half (Half #2) will be devoted to in-class applied/empirical exercises and focused on the material discussed in Half #1. In-class Exercise work will typically not be completed in Half #2, but can be finished over the course of the following week and turned in by/at the start of the next class. Our course TA and I will be assisting students in-class during Half #2.

I will be posting class/exercise material in advance, and will assume that you've had a chance to review that material prior to class. That way we can hit the ground running!

**Course Structure:** There are five graded elements in the course; they are:

### **1. One mid-term exam (30% of course grade)**

There will be one online exam (a Canvas Quiz) covering the empirical methods and applications developed in this course. The exam will be open book, open notes, open universe... with one exception there will be no restrictions on the materials that you can use while taking the exam. That exception is that you are not allowed to collaborate with classmates or other individuals (yes, AI tools such as ChatGPT are "individuals") whilst taking the exam. The exact exam date is yet to be determined, but you'll see that I'm proposing Monday 12/9 (the last day of classes in the Fall semester). Exam grades are curved.

### **2. Weekly exercises (30%)**

As discussed above there will be weekly exercises/assignments/problem-sets devoted to the topic of the week. I will distribute/post these prior to class, so that you are well prepared (there will be some pre-class work). Each exercise will have you working with data and applying the econometrics methods featured in the course. You will be completing Answer Sheets for each exercise, which will be graded on a 10 point scale. Exercises are graded as team assignments. I will not give separate grades for team members, irrespective of how much each teammate might have contributed to the assignment. My hope and expectation is that teams will generally do quite well on these exercises.

As mentioned above, you'll have the rest of the week to complete any exercise work not finished in-class during Half #2. (You will almost always need that additional time to complete the exercises.) At the latest, exercise Answer Sheets should be submitted by the start of the next class, when I will also quickly review the answers to the exercise.

---

<sup>1</sup> [https://www.bc.edu/content/bc-web/academics/sites/university-catalog/policies-procedures.html#academic\\_integrity\\_policies](https://www.bc.edu/content/bc-web/academics/sites/university-catalog/policies-procedures.html#academic_integrity_policies)

## econ3370F24: Topics in Applied Econometrics

These will typically be team assignments (usually with two students per team). I will assign the teams, which will change from exercise to exercise.

### 3. Empirical research project and presentations (27.5%)

This will be a team assignment, which will kick-off with team assignments by Mon Sep 30<sup>th</sup>, after the first five classes. I will assign teams, each of which will have two randomly selected students. Students who would like to work together should let me know their preference by 5PM Friday Sep 27<sup>th</sup>. I will do my best to honor those requests, but I cannot make any teammate guarantees. More details about the actual assignment below.

### 4. wk9 NCAA Football Forecast Challenge (7.5%)

One of the great things about forecasting models is that you get to see how well your model performs. With the wk9 NCAA FFC (*Football Forecast Challenge*) you'll be predicting the outcomes of about 55 NCAA FBS-FBS games played over the course of the weekend beginning Friday Oct 25<sup>th</sup> @ 7:30 PM with the kickoff to BC's *Homecoming* game against Louisville @ Alumni Stadium. This will be a team assignment with two students per team (as usual, I will assign teams... by Weds Oct 9<sup>th</sup>). I will also participate in the *wk9FFC*. Your grade will depend on how well you do relative to the other forecasts (including my forecasts). Stay tuned for more details.

Warning: A general rule with forecasting models is that *less is more...* or put differently, *parsimony is preferred* (overfitting the data can lead to very bad forecasts). And so knowledge of NCAA football may in fact be more of a hindrance than a help.

### 5. Review Quiz (5%)

There will be an online Quiz covering the review material in the first two classes. This is designed to ensure that everyone is up to speed with their 2228 metrics material, and well prepared to tackle the more challenging topics in this course. The Quiz date is yet to be determined, but you'll see below that I am proposing Mon Sep 16<sup>th</sup>. As with the Mid Term exam, this Quiz will be open book, open notes, open universe... but closed colleagues/AI.

**Hardware (for in-class work):** Given the unusual structure of the course, students should bring their laptops with them to class. If you are not able to do this, let me know in advance and I'll arrange for a laptop for you. Note that loaner laptops are available through O'Neill Library.

**Software:** This course will make extensive use of both Excel and Stata.

*Stata:* You should have worked with Stata in your Econometrics course. At the start of the semester, we will review how to access and run Stata through BC's apps server. To avoid traffic jams on the apps server, you may want to purchase a six-month Stata/BE license for \$48. For details, go to <https://www.stata.com/order/new/edu/profplus/student-pricing/>.

*Excel:* This course also makes extensive use of Excel, particularly its *Solver* functionality. You should not take this course if you do not have strong Excel skills. To brush up on your Excel skills, you could take a look at any one of about a gazillion websites. Here's one that students have in the past found useful: <https://www.excel-easy.com/>

**Important:** Unfortunately, the features offered by Excel differ somewhat across platforms and over time. For this course, you will need to have access to Excel's *Analysis ToolPak* and

## econ3370F24: Topics in Applied Econometrics

**SolverAdd-in.** If your version of Excel is anything other than Excel 2011 for the Mac, you should be in great shape, though you may have to install the *Tool-Pak* and *Solver* functionality (that should be straightforward).

If you have Excel 2011 for the Mac, then you have at least two options:

- 1) run Excel on BC's apps server, which has these capabilities and is faster than you might imagine (to login go to <https://bcapps.bc.edu/logon/LogonPoint/index.html> ), or
- 2) go to <https://www.bc.edu/content/bc-web/offices/its/support/software-hardware/software-downloads.html> to see what other solutions BC offers you.

There are other options as well, if you don't like these two:). Let me know and I'll be happy to help.

### The Applied/Empirical Research Project

The empirical research project will kick off with team assignments at the end of September (you'll see that I'm proposing Mon Sep 30<sup>th</sup> ). As mentioned above, there will be two teammates per team. I will randomly assign teams, although... Let me know if you would like to be teamed with a specific teammate, and I will do my best to honor your request... but I cannot make any teammate guarantees. As I said above: *Students who would like to work together should let me know their preference by 5 PM Friday Sep 27<sup>th</sup>.*

There are no restrictions on topics, but your analysis should employ techniques/methods that have been developed/considered in this course. I'm happy to help with topic selection, and more generally over the course of the project.

As you'll see in the Calendar, I plan on setting aside two classes for meetings/presentations in which each team will review their progress to date. This should help to keep projects on schedule and assure a successful outcome.

There are two end-of-semester deliverables: A hardcopy Presentation (PowerPoint or Word or whatever... I just have to be able to print it out) as well as an in-class presentation @ the last class. Your deliverable/presentation should include a detailed discussion of your data, analysis, results and conclusions. As part of that presentation you should discuss what's *new and different* about your work. (You should not fret for one moment over *new and different*... my experience is that once you are working with data, *new and different* appears everywhere you look.)

I'll say more about the project specifics when we kick things off.

At the end of the semester teammates will assess their own and each other's performance using the posted Peer Evaluation form (see posting to Canvas). Students' grades will reflect both their individual performance as well as the quality of the final team submission.

Grading will also reflect degree-of-difficulty. Want my take on degree-of-difficulty? Just ask!

Some of you are working on Senior Theses this semester, and are no doubt wondering if you can use your research project to help with that work. That's fine with me, so long as your partner consents and your Sr Thesis advisor approves. I only have one request: Your research project presentations should make clear what is new and different from your Senior Thesis work. Or put

## **econ3370F24: Topics in Applied Econometrics**

different: Your research projects should be much more than a simple regurgitation of work you've already done on your Senior Thesis.

To review the important research project dates:

- Fri Sep 27<sup>th</sup> : Deadline for letting me know if you have a preferred teammate
- Mon Sep 30<sup>th</sup> : Teams assigned
- Weds Oct 23<sup>rd</sup> : One-on-one meetings to discuss your topics (about six slides per team)
- Weds Nov 13<sup>th</sup> : In-class project update presentations (about six slides per team)
- Weds Dec 4<sup>th</sup>: Research projects due; In-class presentations (about six slides per team)

*Empirical work is slow going... and always slower than expected, even for veterans. Be sure to leave yourself enough time to complete the assignment to your satisfaction.*

**Course Topics:** While the list may change, at the moment I anticipate that we'll be focusing on the following topics over the course of the semester, working with both Stata and Excel:

### **1) Gathering data and building datasets**

### **2) Least Squares Estimation**

- a) Review of Ordinary Least Squares (OLS) estimation and inference, working with Simple and Multiple Linear Regression (SLR and MLR) models (estimation, endogeneity, partial correlations, inference, multicollinearity, heteroskedasticity, dummies, F tests, etc.)
- b) Binary dependent variables I: Linear Probability Models (LPMs)
- c) Functional forms: i) parametric: ln's and exp's, polynomials and splines (linear and cubic); ii) non-parametric: (percentile) dummies and fixed effects, k-nearest neighbor and kernel estimation.
- d) Nonlinear least squares, estimation and inference... including truncated LPMs
- e) Generalized least squares
- f) Resampling and standard errors
- g) Differences-in-Differences (Diff-n-Diff) and Regression Discontinuity Design (RDD) analysis

### **3) Maximum Likelihood Estimation**

- a) Binary dependent variables II: Logit and probit models
- b) More about limited dependent variables: Working with ordered/ordinal data; Ordered logit and probit, and perhaps multinomial logit/probit
- c) Count models (non-negative integer dependent variables): Poisson and negative binomial models

### **4) Choosing covariates**

- a) Machine learning and variable selection algorithms

## econ3370F24: Topics in Applied Econometrics

Topics that we will definitely not, consider in this course: time series (with one small exception); sophisticated panel data techniques; simultaneous equations; Instrumental Variables (IVs); Generalized Method of Moments (GMM) estimation.

**Proposed Schedule of Topics:** The schedule and topics will likely evolve as we work through the semester, but here's a sense of what I have in mind at the moment:

- 1) Weds 8/28 - OLS Review #1: OLS/SLR models (estimation, GOF and inference)
- 2) Weds 9/4 - OLS Review #2: ... and MLR models (estimation, GOF and inference)
- 3) Weds 9/11 – Binary dependent variables (*Linear Probability Models*): ... and functional forms (parametric and non-parametric)

Mon 9/16 - Quiz on Reviews #1 and #2 (exam is after return of graded wk02 Answer Sheets)

- 4) Weds 9/18 - Nonlinear and generalized least squares
- 5) Weds 9/25 - MLE I: Introduction to MLE (*Maximum Likelihood Estimation*): Logit/Probit

Mon 9/30 - Research project teams assigned

- 6) Weds 10/2 - MLE II: MLE continued (working with ordinal data; ordered Logit and Probit)
- 7) Weds 10/9 - *wk9FFC* kickoff: Teams assigned; ratings models (Bradley-Terry (OLS/MLE); points models (OLS); weighted and/or shrinkage estimators)

### ***Fall Break***

- 8) Weds 10/16 - Standard errors (bootstraps and jackknives) and variable selection algos
- 9) Weds 10/23 - One-on-one research project meetings (+/- six slides due Tues 10/22)

Fri 10/25 - *wk9FFC*: forecasts due by 7:30 PM

- 10) Weds 10/30 - MLE III: Estimation and inference; multinomial Logit/Probit
- 11) Weds 11/6 - MLE IV: Count Regression Models (Poisson and Negative Binomial)
- 12) Weds 11/13 - Research project updates: In-class presentations (+/- six slides due Tues 11/12)
- 13) Weds 11/20 - Diff-n-Diff and Regression Discontinuity Design (RDD) Models

### ***Thanksgiving Break***

14) Weds 12/4 - Research projects due; In-class presentations (+/- six slides)

Mon 12/9 - Mid-Term exam (online Canvas quiz)