York University
AP/Econ 4210 3.0 M: Econometrics
Winter 2020

– Syllabus –

Official description

The objective of this course is to provide students with a number of tools for conducting and understanding empirical research in economics. The focus of the course will be on empirical methods for cross-sectional and panel data. The main topic of the course is the linear regression model, its estimation and inference. We cover least squares and maximum likelihood estimators. Other topics include heteroskedasticity, endogeneity, instrumental variable estimation, and simultaneous equations, difference-in-differences. In addition to analytic exercises, students will receive practical questions requiring handling and analyzing data using the statistical software package `R`. Prerequisite: AP/ECON 3500 3.00 or 3210 3.00. Course credit exclusions: GL/ECON 4260 3.00, SC/MATH 3330 3.00. Note: This course can be taken along side Econ 4140.

Note: Econ 3500 is **not** required for this course, provided you've taken Econ 3210. It is highly recommended you have at least **one** of Econ 3210 or Econ 3500.

What we're gonna do

This is an intermediate-level course in econometrics for students at the York (Keele campus). The goal is to equip you with a modern approach to data analysis and econometrics, focusing on the use of data to answer causal questions. You will learn about different empirical techniques that economists use to do so: random assignment, linear regression, difference-in-differences, instrumental variables and regression discontinuity design. The workhouse of the course is the linear regression model, its estimation and inference.

You will learn about applications of these techniques. In particular, you will get hands on experience and gain familiarity with `R` and `Rmarkdown` to get data, manipulate data, perform data visualization, estimate econometric models, and present and communicate results.

Learning Outcomes

By the end of this course you should be able to:

1. Understand the notion of causality, its importance in empirical work, and differentiate causality from correlation or association.
2. Identify the appropriate empirical strategy to answer causal questions using data: random assignment, regressions, instrumental variables, difference-in-differences, and regression discontinuity design, and their associated regression specification.
3. Clearly articulate each method's requirements/assumptions, typical use, and limitations, and know how to interpret their quantitative results.
4. Interpret and comment on tables of estimated coefficients from a wide range of econometric models, in various formats.
5. Use your understanding of the methods to assess the validity and quality of empirical studies, including the ability to judge whether a method may or may not work in a specific research context.
6. Apply these methods to actual data-sets, using the `R` programming language and output and communicate the results in `Rmarkdown`.

Why Econ 4210?

1. Answering questions using data is a big deal:

1. Netflix job posting keywords: Causal Inference; Experimental Design; Advertising Effectiveness,
2. Facebook's Data Science Team features RCTs and causal inference,
3. Amazon offers a reduced form/causal/program evaluation *track*,
4. Google seeks experience in experimental design and causal inference.

From a random Deliveroo job ad (source) looking for an economist:

> Experimenting at this scale presents some unique challenges and we're investing heavily in building a world-class platform for designing, deploying, and analyzing product experiments. We're looking for experts in statistical inference and estimation to join our growing team of data scientists and help us develop innovative statistical solutions for industrial-scale experimentation.

And its not just the private sector – policy evaluation is an important part of government decision making at all levels. **Answering causal questions using data is what this course is about.**

2. You plan on getting a graduate degree

This course, along with the other core economics courses, are essential for success in MA Economics programs. This course, in particular, is a strong signal of a students potential in an MA Economics program and, thus, highly useful in making admittance decisions. If you are planning on an MBA instead, this course is still a great fit, a strong signal of potential, and will help set you apart from other candidates a bit. In Master's of Public Policy and

Master's of Public Administration programs, this course will help you get a leg up on their statistical and program evaluation courses.

3. You want to stand out

There are just under 2500 economics majors at York (Keele) and only about 25 students take this course per year (roughly, just over 1 percent). Despite what Bernie says, you want to be in this **1 percent.**

Why `R`?

1. Its free

`R` is a free, open-source programming language for statistical computing. All of our work in this class will be done using `R`. You will need regular, reliable access to a computer running an up-to-date version of `R`. Download here

`RStudio` is a free, open-source, industry leading, `R` programming environment. It contains a built-in code editor, many features to make working with `R` easier, and works the same way across different operating systems. Use of `RStudio` is required for the labs, and strongly recommended in general. Download here

`R markdown` provides an authoring framework for data science. You can use a single `R Markdown` file to both:
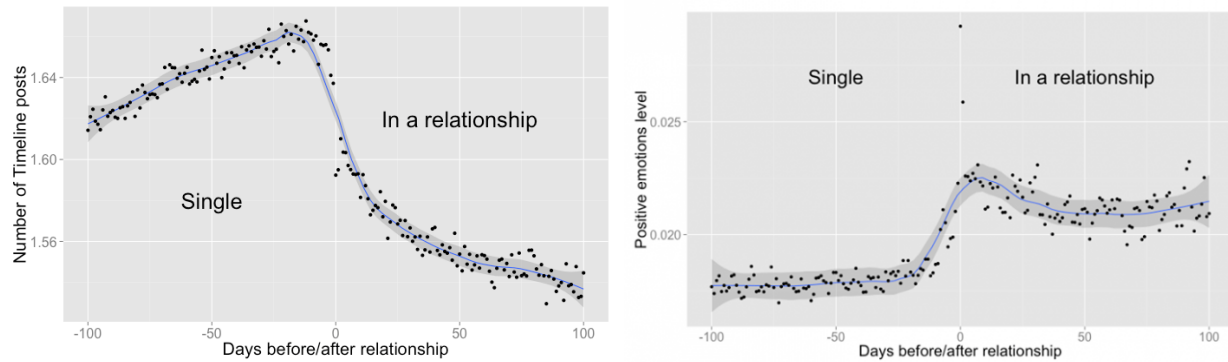
1. save and execute code,
2. generate high quality reports that can be shared with an audience.

2. You'd like a job at some point, and `R` can help.

`R` is used in nearly every industry, particularly in fields that require data analytics. Many large companies use `R` (source):
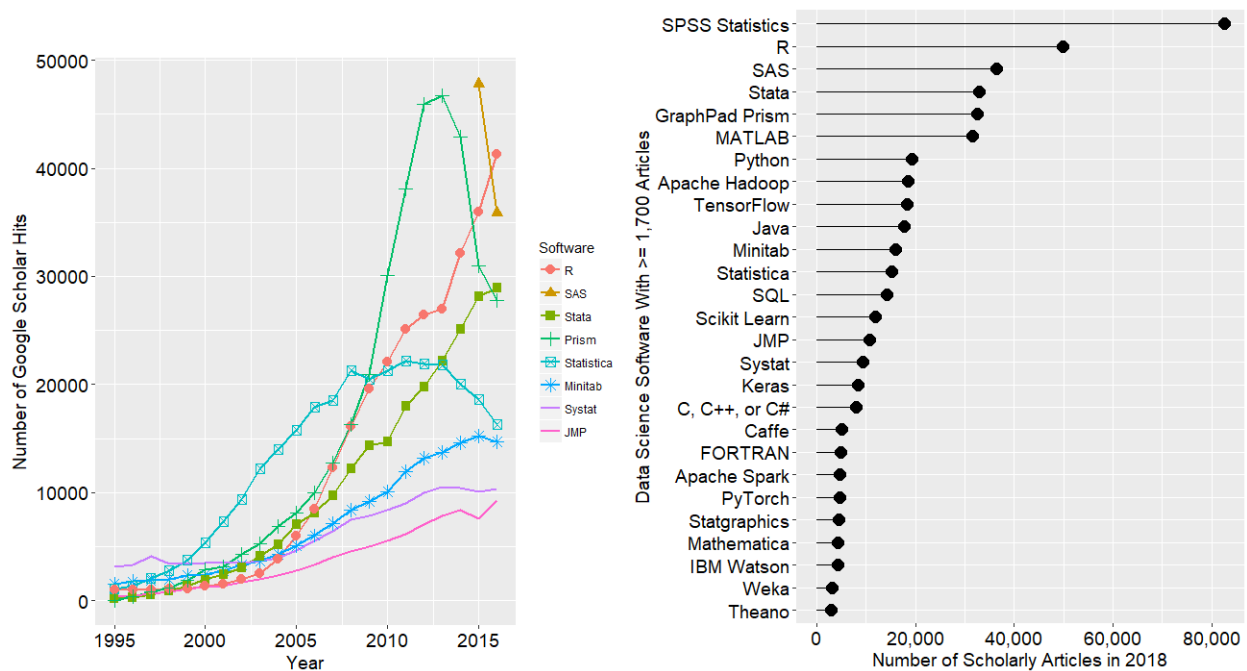
At facebook, `R` is used for data analytics and visualization, among other things (source):

3. `R` is popular and growing in cutting-edge research

`R`'s popularity is growing in scholarly research (source).

Don't take it from me, here are some other sources:

1. Fast Comapny,
2. Turn a hobby into a career,
3. 8 Reasons,
4. Because these guys use `R`, and they do great data visualization.

Lectures and tutorials

The course is built on weekly 2-hour lectures, followed by one-hour tutorials, but some weeks will feature a 3-hour lecture. See the Course schedule for more details.

Lectures

Before each lecture, you are expected to do the required readings. Readings will usually be taken from the textbook, but will occasionally come from additional sources, typically research papers. Two-hour lectures usually have a 10min break in the middle.

Tutorials

In the tutorials, we deal with the practical application of the concepts and methods used in class. This will typically involve R. Bring a laptop. Or make a friend with someone who owns a laptop. Or do both.

How to be successful in this course.

Eighty percent of success is showing up – **Woody Allen**

There is a recent trend among students to not attend classes. This, in general, is not a good idea. The most straightforward way to success (ie, a good grade and perhaps a letter of reference for a grad program) is to show up, pay attention, participate, do the assignments (they are easy, low-stress grade inflators), and not leave all of your studying to the day before the quiz/final. I want you to succeed. I am willing to help if you struggle. But there has to be a certain level of commitment on your end, too.

Comportment

Are you a strategically adept effort-minimizing, goal-seeking, time-manager in the grand York tradition? If so, this class is probably not for you. Econometric success requires a certain level of commitment. Like most things worth doing, econometrics also requires focus and attention. In this spirit, I ask you not to bring food to class and, during lectures, to leave electronic and other toys shut off and put away (this prohibition includes, but is not limited to: laptops and tablets; ipods; phones; Wii, Xbox, or Playstation consoles; VR headsets; Google Glass; any new invention that I am not yet aware of, but that I am likely too old to understand).

	Day	Time	Duration	Place
Lecture:	T	16:00	180	BC 215
Office Hours (Ben):	T	14:00	60	1076 Vari Hall

Lectures and Office Hours

Office Hours

I will be available for office hours. During these times students are encouraged to drop by. If you cannot attend the scheduled office hours, I will be available by appointment as well. I will provide additional office hours before the midterm and final exam.

E-mail

E-mail is a good way to contact me. I will attempt to answer reasonable questions via email as quickly as possible. Please include "Econ 4210" in the subject line to facilitate this process. Your name (as it is written on Moodle) and student ID are required in every email if you expect to get a response. My answering-your-email-vibe depends on the quality of your email. Here are some examples of actual emails:

Bad emails

```
hi
could you plx tell me your office hour for thurs? and r u going upload
solution of sample papr?
```

```
Dear Sir:
Could we have the solution of the questions?
```

```
Hello Sir
My final calculated grade is 47 and I have missed two assignments, is it
possible that can you please allow me to submit the assignments now so
that I can pass the course.
```

My response

```r
for (i in 1:3) {
  print(paste0("nope."))
}
```

```
## [1] "nope."
## [1] "nope."
## [1] "nope."
```

In every case, the student should have known the answer to their question by listening in class and/or reading the syllabus.

Course Work

Tutorial Questions

During the course of the term, there will be five (5) in-class tutorial questions. These can be done in groups of two (2), but each student must upload their own tutorial question. Each Tutorial question is worth 5 percent of your grade. Only the best four (4) Tutorial questions will count toward your final grade. Tutorial questions will be of applied nature, done in `R` and `R markdown`. If a student finds a problem with the grading of a Tutorial question, she/he should immediately talk to the me. The deadline for a regrading request is **one week** from the day the grade is posted on Moodle , regardless of when the student actually receives it.

Homework Assignments

During the course of the term, there will be four (4) homework assignments to be done in `R` and `R markdown`. Each homework assignment will be worth 5 percent of your final grade. This work can be done in groups up to four, but each student must upload their own assignment. No work will be accepted after the due date. Students are encouraged to study the suggested answers to each problem set, regardless of their performances on the problem set. If a student finds a problem with the grading of a problem set, she/he should immediately talk to the me. The deadline for a regrading request is **one week** from the day the problem set is returned to the class, regardless of when the student actually receives it.

Homework assignments dates and instructions are posted well in advance, and it is your responsibility to ensure adequate time to complete the work and deal with any issues, including technical issues. Failure to submit an assignment on time will result in a grade of zero. Assignments are considered submitted by the time all the files have been uploaded in the correct format to Moodle according to the assignment instruction. Assignments cannot be submitted late. There will be no exceptions. Make sure to allow ample time for submission before the deadline; excuses such as: *the website is slow*, *I only submitted one minute late*, *I have had a stomach bug on the last day before deadline*, *I forgot to upload one of the files*, etc. are not valid excuses. These rules are there to limit unwarranted individual requests, which take up valuable time that I could spend improving the course content.

Quizzes

During the course of the term, there will be five (5) small quizzes. The quizzes will be done individually. The format of the quizzes will be similar in style to the type of questions you will get on the final exam. Only the best four (4) quizzes will count toward your final grade. Each quiz will be approximately 30 minutes long. The dates of the quizzes and the topics

covered will be announced in class and posted on Moodle prior to the quiz. The quizzes are not meant to be difficult; rather, they are intended to make sure each student is following the material on a regular basis and to identify students who are struggling and who might require additional help. You can think of them as a midterm, just spread out over the term.

Missed Tutorial Questions and Quizzes

There are five (5) of each, and only the best four (4) count toward your final grade. If you miss a quiz or a tutorial question, it automatically becomes your lowest grade and, thus, there will be no retake. If you miss one Tutorial question or quiz, I do not require any documentation, email, or other form of communication. I will simply assess a zero for that component. If you miss more than one Tutorial or Quiz, then I require an Attending Physician statement **within** one week of the missed component. The weight of that component will be shifted toward the final exam.

A student may write all five (5) quizzes and tutorial questions regardless of previous performance. However, only the best four (4) will count toward the final grade. Tutorial questions and quizzes are not interchangeable.

Final exam

A final exam, scheduled by the registrar, will be worth 40 percent of the final grade. It will be a two-hour exam, consisting of two 1-hour written components, written with a 3-hour time block. I will discuss the details at the appropriate time. I will post a practice final exam so that students are fully aware of the style of questions that will appear on the final exam. I will provide solutions to the pratice final exam. You do not have to ask for them, I promise I will post them. Emailing me will not speed up this process. The final exam is cumulative.

Grade

The final grade in the course will be determined as follows, with the following exception below:

Course Work	Number	Weight
Homework	4	20%
Tutorials	4	20%
Quizzes	4	20%
Final Exam	1	40%

Since the homework and the quizzes represent a (potential) group effort, a passing grade on the weighted average of the quizzes and final exam are required to pass the course, regardless of the student's overall grade. Ie., to pass, it must be the case that:

$$(.4 \times \text{Final Exam Grade} + .2 \times \text{Total Quiz Grade}) \cdot \frac{1}{.6} > 50 \text{ percent}$$

The grading scheme for the course conforms to the 9-point grading system used in under-graduate programs at York (e.g., A+ = 9, A = 8, B+ = 7, C+ = 5, etc.). Assignments and tests will bear either a letter grade designation or a corresponding number grade (e.g. A+ = 90 to 100, A = 80 to 90, B+ = 75 to 79, etc.).

Why is there so much course work?

Because I want everyone to do well. It helps to have consistent practice and feedback. The homework component is not overly demanding or time consuming. The tutorial and quizzes are done in class time and meant to emphasize key concepts and build your skill-set through practice and feedback. Since each new concept builds on the previous ones, its important to keep up with the material and have a chance to demonstrate your mastery of each concept on a routine basis. This course outline attempts to establish a feedback loop: Students complete a task and are evaluated, receive feedback, and can adjust accordingly. The feedback loop is a process to check and confirm the understanding of specific, manageable, focused concepts at regular intervals to build capacity and depth over time. The feedback is important for me, also, so that I can address areas of weakness. And it works.

Reading

The following textbook is required for the course:

Wooldridge, Jeffery M., *Introductory Econometrics: A Modern Approach 5th Edition.* Cengage Learning.

Topics to be Covered

1. Review of Basic Statistical Concepts

2. Review of Regression and Inference

3. Multiple Regression Analysis: OLS Asymptotics

4. Multiple Regression Analysis: Additional topics and issues

5. Instrumental Variables Estimation

6. Introduction to Panel Estimation

7. Limited Dependent Variables

8. Program Evaluation

Important Information

All students are expected to familiarize themselves with the following information, available on the Senate Committee on Curriculum & Academic Standards webpage; http://www.yorku.ca/secretariat/policies/index-policies.html/

- York's Academic Honesty Policy and Procedures/Academic Integrity Website

- Ethics Review Process for research involving human participants

- Course requirement accommodation for students with disabilities, including physical, medical, systemic, learning and psychiatric disabilities

- Student Conduct Standards

- Religious Observance Accommodation

The departments guidelines for deferred standing

Can be found at http://dept.econ.yorku.ca/undergraduate/Deferred/_Standing1.pdf