

York University | Winter 2023 | AP/Econ 4210 3.0 M: Econometrics

– 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.

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: random assignment, linear regression, difference-in-differences, and instrumental variables. The workhorse 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.

Lectures and Office Hours

	Day	Time	Duration (min)	Place
Lecture:	T	14:30	~ 60-80	ACW 106
Tutorial:	T	16:15	< 60	ACW 106
Office Hours (Ben):	F	2:00	~ 60	Zoom

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 via reading/video that I will create and post.

Tutorials

In the tutorials, we deal with the practical application of the concepts and methods used in class. This will involve R and Rstudio. You must have a working computer, a current installation of R and Rstudio and an internet connection.

Please bring your laptop to class so that you can fully benefit from the tutorials

Office Hours

I will be available for office hours. During these times students are encouraged to drop into a zoom session. If you cannot attend the scheduled office hours, I will be available by appointment as well.

E-mail and the course message board

E-mail is not a good way to contact me for questions regarding the course. For these questions, please use the e-class message board. I will attempt to answer all questions posted there within a day or so. However, I will not answer questions about course material via email and will simply not respond.

Examples:

Q: Professor, I keep getting this error in my code: Error in evaluating the argument 'object' in selecting a method for function 'summary': object 'wage' not found. What should I do?

A: You should post on the message board.

Q: Professor, I can't figure out MyLabs question 6.5. Can you help?

A: Sure, if you post on the message board.

These answers, of course, are hypothetical – because I wouldn't actually bother to respond.

However, if you are having a **personal issue** and would like to talk to me, then please feel free to email me anytime. We can even meet over zoom if that is easier.

Course Work

The course is broken down into three Modules. **Each** module contains:

1. Quizzes that are done on MyLabs,
2. Practice Exams done in Rstudio and submitted on e-class,
3. A module Exam.

There is no midterm or final exam in the traditional sense. Instead, each module contains an exam, so there are three exam throughout the term. The last one will be scheduled by the registrar, but it will be the same format as the other two and be module specific.

Practice Exams

During the course of the term, there will be three Practice Exams, one for each module, to be done in **R** and **R markdown**. This work can be done in groups up to four, but each student must upload their own Practice Exam. If a student finds a problem with the grading of a Practice Exam, she/he should immediately talk to the me. The deadline for a regrading request is **one week** from the day the Practice Exam grade is posted to e-class, regardless of when the student actually receives it.

Practice Exam due 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*.

- Empirical Practice Exams will consist of longer-form empirical questions and must be completed in **Rstudio** and submitted as an **HTML** document constructed in **Rmarkdown**. Practice Exams may involve downloading and loading data, installing and loading packages, writing code, constructing (nice) tables and figures, interpreting and communicating statistical output.
- The Practice schedule will be posted well in advance. There will be at least 14 days to complete an assignment.
- The Practice Exams are designed to prepare students for the written component of module exams.

Quizzes

During the course of the term, there will be five (5) quizzes. The quizzes will be done individually. Each quiz will be approximately 45 minutes to an hour long.

- Quizzes will consist of answering numeric, multiple choice, and true-false type questions on MyLabs. Some questions will involve data and I recommend you use R to answer these questions.
- The Quiz schedule will be posted well in advance. There will be at least 3 days to complete each quiz. Students will have one attempt per quiz. Each quiz has a 45 minute time limit.
- Quizzes are designed to help students on the non-written part of the module exams.

Module Exams

There will be three (3) exams throughout the semester, one for each module. The first two will be held in class, and the last will be scheduled by the registrar.

Missed coursework

1. **Practice Exams:** Failing to submit an PE will result in a **grade of zero, no exceptions**. I will accept late PEs up to 5 days, with a ten percent deduction from the PE grade for each day late. For example, suppose your score on the PE was 18/20, but submitted the PE 3 days late: you would receive credit for 12.6/20.

There is no need to contact me for late/missed PE, and I will not respond to such emails.

2. **Quizzes:** Failing to submit quiz will result in a **grade of zero, no exceptions**. However, if you are ill for the entire quiz period, you can defer your quiz. **All deferred quizzes will be offered on April 8th**. This is non-negotiable. I will not offer alternative, personalized times for each student. What if you defer more than one quiz? Then April 8th will be a long day for you.

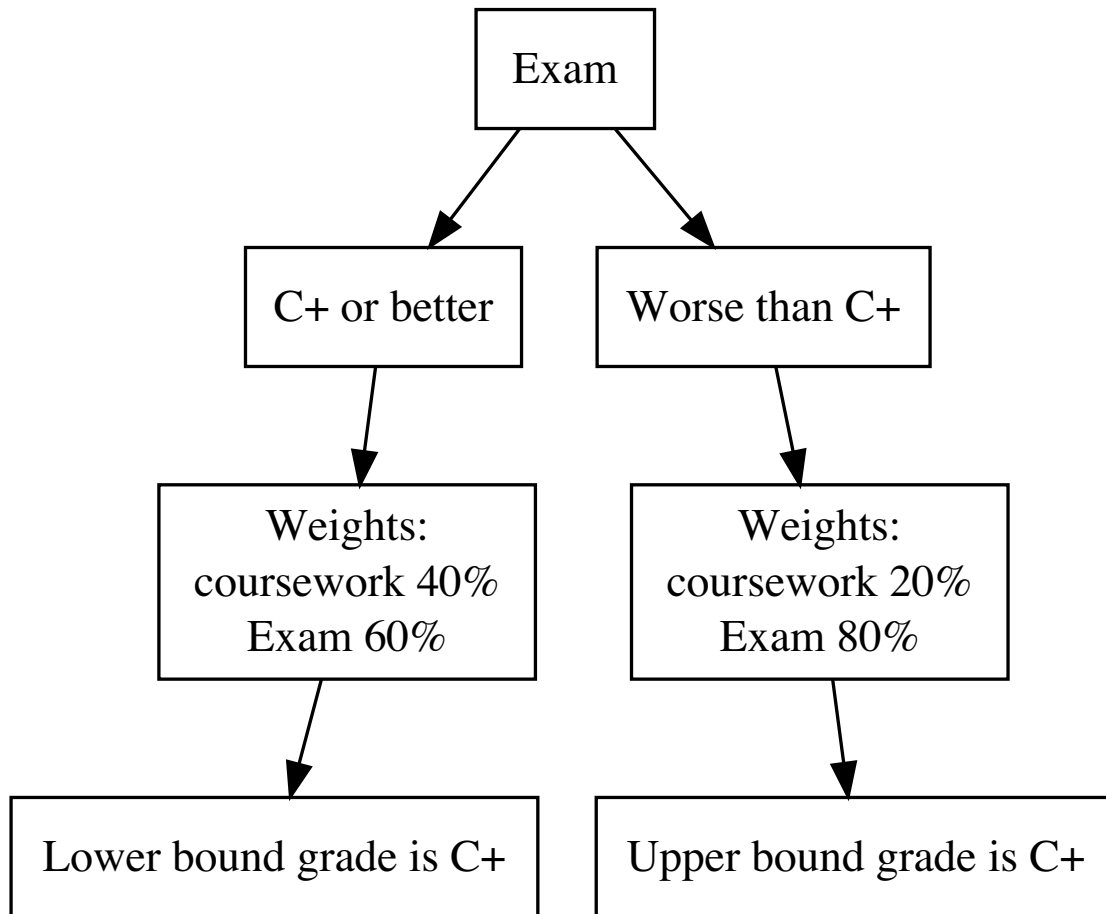
If you wish to defer a quiz, you must email me *before* the quiz.

3. **Module Exam:** Students missing the module exam must provide detailed documentation in the manner required by the Faculty of Liberal Arts and Professional Studies and the Department of Economics. You can find the requirements on the following link: [Handbook](#). If appropriate documentation is provided, there will be a **deferred module exam offered on the day of the final exam**.

How does this work? I will book a 3-hour time slot for the final exam. The Module C exam is 1.5 hours. Students writing only the Module C exam can then leave. Students writing a deferred exam will stay and complete their deferred module exam. Students deferring Module C or both Module A and B exams will **write a deferred exam in late-July**.

Grading scheme

The course contains three modules. Each module has *coursework* and a *module exam*. The coursework consists of Quizzes and Practice Exams. **The weights of each component depend on your performance in the exam**. In particular, the flow diagram below indicates that for grades C+ or better on the Exam, the weight of the coursework is 40 percent. Below this performance threshold, the weight of the coursework shrinks to 20 percent each, meaning the weight of the exam itself increases. In addition to this, there is a lower bound grade of C+ for students obtaining a C+ or better on the Exam. In this case, poor performance on Quizzes and Assignments will have a negative impact on your grade, but it is limited. On the other hand, for students obtaining a grade worse than a C+ on the Exam, the highest grade that can be obtained in the module is a C+. Thus, Quizzes and Assignments can improve your grade, but it is limited.



Thus, the formula determining your module-grade is as follows:

$$\text{Exam} \begin{cases} \geq 65 & \text{grade} = \max(0.60 \cdot \text{Exam} + 0.40 \cdot \text{Coursework}, 65) \\ < 65 & \text{grade} = \min(0.80 \cdot \text{Exam} + 0.20 \cdot \text{Coursework}, 65) \end{cases}$$

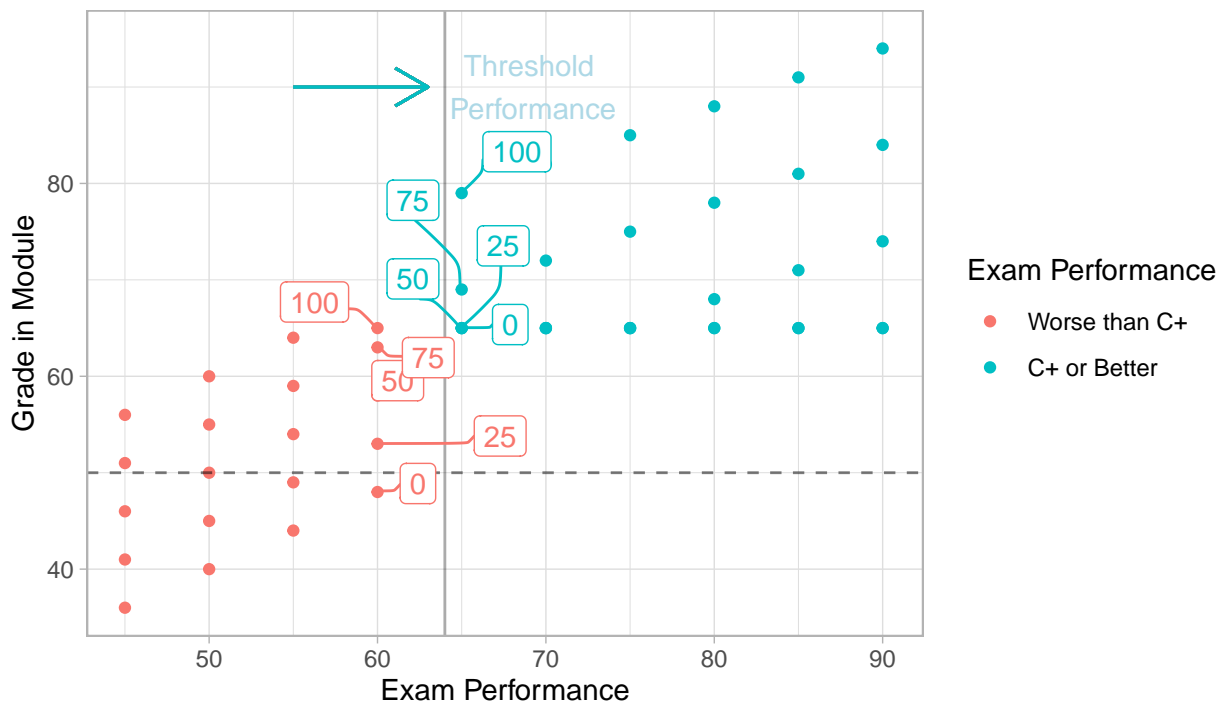
Below is a visualization of this formula. On the x -axis are hypothetical grades on the module exam. On the y -axis are the corresponding module grade. Each dot represents a student with 5 different hypothetical performances on the coursework (average of assignment and quizzes). The teal dots represent students who've obtained a C+ or better on the module exam. The red dots represent students who have obtained below C+ on the module exam. The labelled points show potential different performances in coursework. Note the **asymmetry** around the performance threshold. In particular, take two students:

1. The first student gets 65 percent on the exam, and zero on coursework. This student receives a **C+** in the module.
2. The 2nd student gets 64 percent on the exam, and zero on coursework. This student receives a **D** in the module.

However, picture two other students:

1. The first student gets 65 percent on the exam, and 100 percent on coursework. This student receives an **B+** in the module.
2. The 2nd student gets 64 percent on the exam, and 100 percent on coursework. This student receives a **C+** in the module.

Take away: Past the performance threshold, coursework can help your grade a lot, but not below it. On the other hand, the impact of poor performance on coursework on your grade is limited above the performance threshold, but not below it.



The labels show average coursework grades in five different scenarios. To the left of the performance threshold, the coursework grades do not help your final grade much, and the maximum grade that can be obtained is a C+ (60 percent). To the right of the performance threshold, assignment and quiz grades can help your final score, but cannot hurt it much (lower bound is a C+).

Final grade

Final grades are determined via a weighted average of performance in each module. The weights are as follows:

Table 2: Module weights

Module	Weight
A	.35
B	.40
C	.25

The grading scheme for the course conforms to the 9-point grading system used in undergraduate programs at York (e.g., A+ = 9, A = 8, B+ = 7, C+ = 5, etc.).

Important dates

1. **Drop deadline:** Last date to drop a course without receiving a grade: March 11
2. **Course Withdrawal Period** (withdraw from a course and receive a grade of “W” on transcript): March 12 - April 8

Reading

The following textbook is required for the course:

Stock and Watson, *Introduction to Econometrics, 4th Edition*. Pearson. With MyLabs.

Topics to be Covered

- Introduction to R, Rstudio and R markdown
- Basic data analysis/programming/data visualization
- Review of Basic Statistical Concepts
- Review of Regression and Inference
- Multiple Regression Analysis: Additional topics and issues
- Instrumental Variables Estimation
- Introduction to Panel Estimation/ difference-in-differences
- Limited Dependent Variables
- Program Evaluation/potential outcomes
- Introduction to Machine Learning/Prediction

Important Information

All students are expected to familiarize themselves with the following information, available on the Senate Committee on Curriculum & Academic Standards [webpage](#):

- 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 [this link](#)

Why Econ 4210?

1. Answering questions using data is a big deal:

In both the public and private sectors, decision/policy makers are increasingly relying on data to make informed, data-based decisions. As an example, there are some characteristics of several leading tech firms' job postings:

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 student's 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). 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:

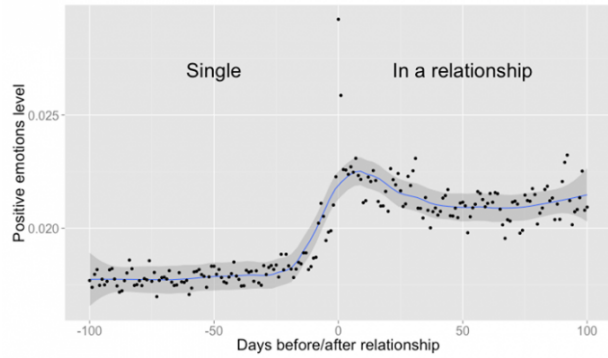
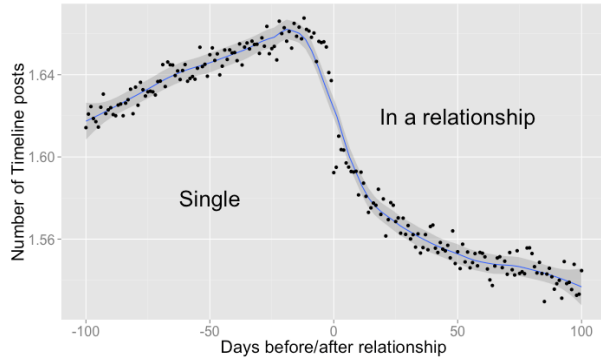
1. save and execute code,
2. generate high quality reports that can be shared with an audience.

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](#)):

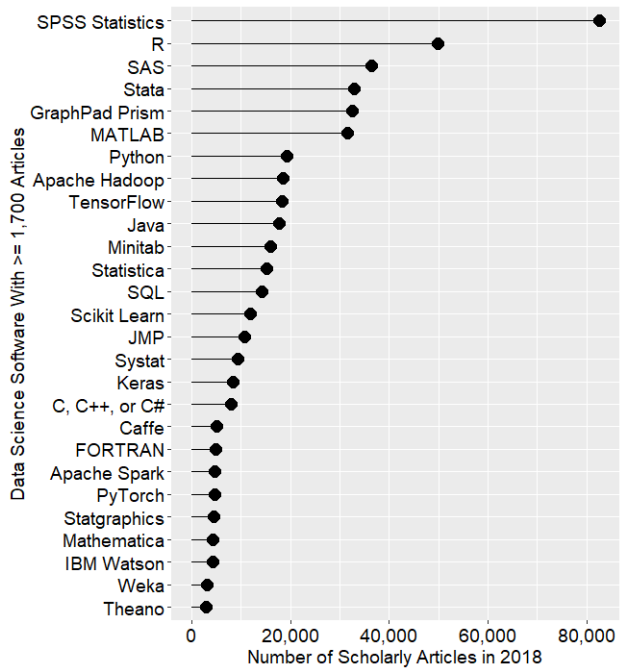
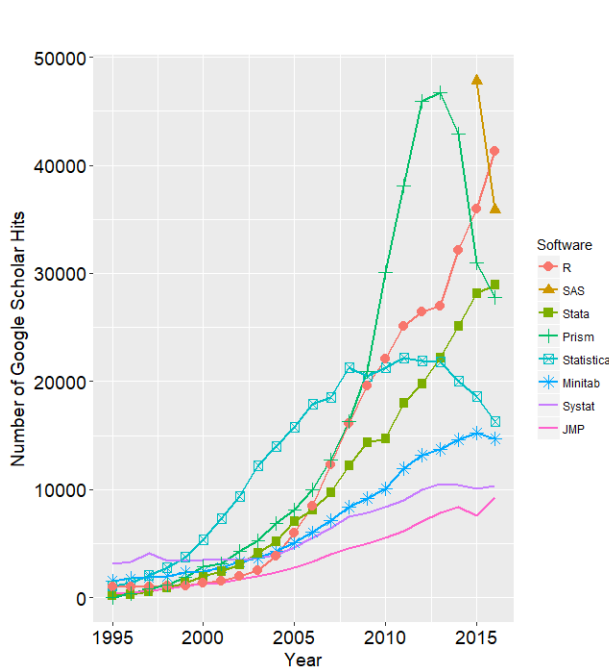


At facebook, R is used for data analytics and visualization, among other things ([source](#)):



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 Company](#),
2. [Turn a hobby into a career](#),
3. [8 Reasons](#),
4. [Because these guys use R, and they do great data visualization](#).