York University
Fall 2026
AP/ECON 4210 3.0
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 alongside Econ 4140.

Econ 3500 is **not** required, provided you have taken Econ 3210. Have at least one of them.

This is the second course in the econometrics sequence. Econ 3210 asks one question, **does X cause Y, and how would we know**, and answers it with randomized experiments and regression. This course starts by re-doing that core quickly, so that everyone is on the same footing whichever section of 3210 they took, and then goes where 3210 does not: what to do when you cannot run the experiment and cannot control for the right things. Difference-in-differences, instrumental variables, and regression discontinuity are the three answers economists use most.

The course also takes regression seriously as a **predictive** tool, which is a different job from estimating a causal effect, and ends with your first look at machine learning: sample splitting, penalized regression, and trees.

R is hands-on in this course. You install it, write it in every class, and are examined on it under supervision. The vocabulary is deliberately small, and everything you need arrives in three lab packs, the first in week 1.

Course Objectives

By the end of this course, you will be able to:

1. **Explain and apply the potential-outcomes framework** — state a causal question precisely, decompose a comparison into a causal effect and selection bias, and explain what randomization does.
2. **Read and estimate regressions for two different purposes** — as a controlled comparison whose causal interpretation rests on the conditional independence assumption, and as a predictor, and say which use is in play.

3. **Use the three quasi-experimental designs** — difference-in-differences, instrumental variables, and regression discontinuity: state the identifying assumption, estimate the design in R, and read its output critically.
4. **Build and evaluate a predictive model** — split a sample, measure out-of-sample error, use cross-validation, and fit penalized regressions and tree-based models.
5. **Work in R with a small, reliable toolkit** — `tidyverse` for data, `fixest::feols` for every model, `modelsummary` for every table — and produce a reproducible R Markdown document under time pressure.
6. **Read an empirical paper the way a referee does** — identify the design, the assumption it rests on, and what would break it.

Lectures and Office Hours

	Day	Time	Duration	Place
Lecture and lab	T	14:30 – 17:20	170 min	HNE 030, Keele campus
Office hours (Ben)	TBA	TBA		TBA

How a Tuesday works

Every class has three parts: short written problems on last week's material, a lecture block, and a **lab** in which you apply the day's idea to data in R on your own laptop. On test days the lab is replaced by the test; on lab-exam days the lab *is* the exam.

Ordinary Tuesday		Test Tuesday		Lab-exam Tuesday	
14:30 – 14:55	Problems on last week	14:30 – 15:30	Lecture	14:30 – 14:55	Problems on last week
14:55 – 15:55	Lecture	15:30 – 16:00	Review	14:55 – 15:55	Lecture
15:55 – 16:10	Break	16:00 – 16:15	Break	15:55 – 16:10	Break
16:10 – 17:20	Lab	**16:15 – 17:15**	**Test**	**16:10 – 17:20**	**Lab exam**

Bring your laptop to every class. The lab is not optional and it is where the R happens.

Readings

Read the assigned pages of the textbook **before** class. The Tuesday problems draw on the reading as well as on lecture.

Office Hours

Drop in during the scheduled time, or e-mail for an appointment.

E-mail and the course message board

Use the **eClass message board** for course questions; I aim to reply within one business day. I do not answer content questions by e-mail. For personal or private matters, e-mail is appropriate.

Assessment

Everything that counts toward your grade is done **in class, supervised.** There are no graded take-home assignments or quizzes.

Component	Weight	Dates
Three written tests, **best two count**, 20% each	40%	Oct 6, Nov 3, Nov 17
Two lab exams in R, **best one counts**	20%	Oct 27, Nov 24
Final exam	40%	Registrar-scheduled, December

You must pass the final exam to pass the course.

The written tests

- One hour, in class, the last hour of the Tuesday listed. Closed book. A one-page formula sheet is provided. Non-programmable calculators permitted.
- Each test covers the sessions since the previous test. Not cumulative, but the material builds.
- Every question has the same shape: a table, a figure, or a short description of a study, then compute a number, choose the correct interpretation, or explain in a sentence. Some questions show R output or a few lines of R code and ask what they do.
- Grades are returned the following Tuesday.

The lab exams

- Seventy minutes, in the lab slot, on your own laptop, using the course R project. You are given a dataset and an R Markdown file with the structure in place; you complete the marked lines and interpret the output, exactly as in the weekly labs.
- You submit the `.Rmd` and the knitted `.html` to eClass before the bell. Supervised: no messaging, no generative AI, no collaboration; internet only for eClass and package documentation.
- Lab exam 1 covers the labs on regression and difference-in-differences. Lab exam 2 covers the causal toolkit in R: regression, difference-in-differences, instrumental variables, and regression discontinuity. The prediction weeks are assessed on the final.

Best-of rules

Your lowest written test is dropped, and your lower lab exam is dropped, automatically. Because of this, **there are no makeup tests or makeup lab exams.** If you miss one for any reason, that is your dropped one. No documentation is needed and none is accepted. If you miss two written tests, the second counts as zero; if you miss both lab exams, the lab component is zero.

The final exam

- Registrar-scheduled, December. Cumulative. Closed book, formula sheet provided, calculators permitted.
- The anchor is a **capstone question**: an unfamiliar published study with its tables, which you read and evaluate, identifying the design and the assumption it rests on. The rest is the same kind of question as the tests, drawn from the whole course, including the prediction weeks.
- **Deferred standing.** If illness or other circumstances prevent you from writing the final, submit the Department of Economics deferred exam request: https://www.yorku.ca/laps/econ/undergraduate-programs/academic-resources/department-policies/deferred-standing/. Deferred standing applies to the final only; tests and lab exams are covered by the best-of rules.

Practice

Practice problems with full solutions are posted for every topic, and every weekly lab is practice for the lab exams. A practice test with solutions is posted before Test 1. None of it is collected or graded.

Why this system?

- **Every mark is earned in the room**, on paper or in R, under supervision.
- **One bad day does not matter**, on either track.
- **R counts, and it is assessed the way it is used**: writing and interpreting code against real data, in the format you practise every week.
- **The final has teeth**: 40%, cumulative, must pass.

Accessibility: Students requiring accommodations should contact Student Accessibility Services and the instructor as early as possible in the term.

Schedule

Table 4: Fall 2026 schedule (subject to change; the eClass version is authoritative). Readings are from Mastering 'Metrics.

#	Date	Lecture	Lab	Reading	Assessment
1	Sep 15	Causal questions. Prediction vs causation. How the course works	Lab 0: R set-up, the course project, first data, first plot	Intro; Ch. 1 to Table 1.1	
2	Sep 22	Chapter 1 in one session: potential outcomes, selection bias, randomization; inference; difference in means	Lab 1: a randomized trial in R	Ch. 1 and appendix	
3	Sep 29	Chapter 2 in one session: regression as controlled comparison and as prediction; OVB; the CIA	Lab 2: the regression toolkit	Ch. 2 and appendix	
4	**Oct 6**	**Functional form for prediction and for interpretation: logs, interactions, polynomials**		**Ch. 6 opening**	**Test 1: sessions 1–3**
	Oct 13	*Reading week — no class*			
5	Oct 20	Difference-in-differences	Lab 3: DiD	Ch. 5	
6	**Oct 27**	**Instrumental variables I: the problem, the instrument, the ratio**		**Ch. 3**	**Lab exam 1: labs 0–3**
7	**Nov 3**	**Instrumental variables II: 2SLS, weak instruments, what IV estimates**		**Ch. 3**	**Test 2: sessions 4–6**
8	Nov 10	Regression discontinuity	Lab 4: IV and RD in R	Ch. 4	
9	**Nov 17**	**Prediction I: in-sample vs out-of-sample, sample splitting, cross-validation**			**Test 3: sessions 7–8**
10	**Nov 24**	**Prediction II: ridge and lasso**			**Lab exam 2: labs 0–4**
11	Dec 1	Prediction III: trees, bagging, random forests	Lab 5: sample splitting, glmnet, trees		
12	Dec 8	Capstone: reading one study two ways. Review			

Important dates. Last day to drop without a grade: **November 10**. Add/drop/withdrawal dates for 2026–27: https://registrar.yorku.ca/enrol/dates/2026-2027/fall-winter. Dec 8 is a class day (confirmed).

Required Course Materials

Textbook (required): Angrist, Joshua D. and Jörn-Steffen Pischke, *Mastering 'Metrics: The Path from Cause to Effect*, Princeton University Press, 2015. **Estimated cost: about $61 CAD** (paperback, Amazon.ca, September 2026). No access code. Any format, used copy, or earlier printing is fine. If you took Econ 3210 in Fall 2026 you already have it; this course uses Chapters 1–2 as review and Chapters 3–6 as new material.

Free companions (optional):

- Cunningham, Scott, *Causal Inference: The Mixtape*, https://mixtape.scunning.com/ — free online; the differences-in-differences and regression-discontinuity chapters go further than Mastering 'Metrics does, with R code.
- Hanck, Arnold, Gerber and Schwardt, *Introduction to Econometrics with R*, https://www.econometrics-with-r.org/ — regression mechanics and R.
- James, Witten, Hastie and Tibshirani, *An Introduction to Statistical Learning with Applications in R*, https://www.statlearning.com/ — free PDF; Chapters 2, 5, 6 and 8 cover the prediction weeks.

Software: R and RStudio, both free. Install both before the first class if you can; Lab 0 walks through it if you cannot. In week 1 you download the **course R project** as **Lab Pack I**: a zip with the project folder, a set-up script that installs the packages, and the data and lab files for sessions 1–3. Two further packs, **Lab Pack II** (DiD, IV, RD; posted by Oct 20) and **Lab Pack III** (prediction; posted by Dec 1), add files to the same folder. Everything in the course runs inside that one project.

Formula sheet: the one-page sheet provided with every test and the final is posted on eClass from week 1.

Technical Requirements

A laptop that runs R and RStudio, brought to every class, and eClass. That is all. There is no online homework platform.

Academic Integrity

Why: Upholding high standards of Academic Integrity is essential for successful academic discourse, and to the pursuit of learning and scholarship. Respecting and enforcing these standards at our university ensures that the degree you earn is a strong signal of your learning and academic achievement.

How: You uphold high standards of Academic Integrity by:

- doing and submitting only your own work
- not permitting others to use your answers
- giving credit to others for their ideas and works, including the use of generative AI
- quoting your sources, correctly paraphrasing and summarizing

In this course all graded work is done in class under supervision. For written tests the rules are the usual exam rules. For lab exams: your own laptop, the course project, no messaging or generative AI, no collaboration, and submission on eClass before the bell. Multiple versions of each test are used.

More information: If you have questions or concerns about what constitutes appropriate academic behaviour, please consult with me or seek additional resources.

Enforcement: York University takes its responsibility to uphold high academic standards very seriously. The Senate Policy on Academic Conduct outlines behaviours that constitute academic dishonesty and processes for addressing academic offences. Sanctions in cases of cheating can include temporary and permanent records on your transcript, suspension from the course, or suspension from the University.

It is not up to the instructor: If I observe behaviour that might involve academic dishonesty, I am required to report it. It is not up to me to judge whether the behaviour was unintended, a minor infraction, a first-time infraction, etc. I am professionally bound to report any suspicious behaviour.

Why Econ 4210?

Answering questions with data is a big deal. In both the public and private sectors, decisions increasingly rest on data. Job postings at Netflix, Meta, Amazon and Google ask for causal inference and experimental design by name; Amazon runs a program-evaluation track for economists. Governments evaluate policy the same way. Answering causal questions with data is what this course is about.

You plan on a graduate degree. This course, with the other core economics courses, is essential for MA Economics programs and a strong signal in admissions. It transfers directly to MBA, MPP and MPA programs, where it puts you ahead in the statistics and program-evaluation courses.

You want to stand out. There are close to 2,500 economics majors at York and about 25 take this course each year.

Why R?

It is free, it is what economists and data scientists use, and it is what you will be asked about in interviews. R Markdown lets one file hold your code and your write-up, which is how the labs and the lab exams work. This course keeps the R vocabulary deliberately small so that the tool never gets in the way of the ideas: a handful of `tidyverse` verbs, `ggplot`, `feols`, and `modelsummary` carry the whole term.