Colgate University
ECON 375: Applied Econometrics[†]
Spring 2022 (1/24 - 5/13)

Professor:	Pierce Donovan	pdonovan@colgate.edu
Lectures:	Persson 209	TR 8:30 - 9:45 am (A)
		TR 9:55 - 11:10 am (B)
Labs:	Persson 209	F 8:00 - 10:00 am (A)
		F 10:30 - 12:30 pm (B)
Office Hours:	Spear 9	MW 10:15 - 11:30 am

Course Description and Objectives

Statistics provides us with a set of tools to understand variation and uncertainty, which are generally things humans try to avoid. But if we learn about some of the ways to collect, analyze, and draw conclusions from data—which on their own emit an inherently *noisy* signal—we can begin to recover some knowledge, or, more precisely, recognize patterns that may not be due to chance.

Econometrics builds on this by taking the interpretation and presentation of statistical analyses particularly seriously. Good econometric work is different from other data sciences because it involves *clear, causal thinking*. When it comes to understanding the relationships between different variables, many questions worth asking require careful and creative analysis. Do smaller class sizes promote better student performance? Can city air quality be improved by implementing driving restrictions? How does increased police presence affect crime rates? During this course you will begin to understand how to answer questions like these in a convincing manner.

In this course, your task is twofold: (1) learn how to assess the strengths and weaknesses of different empirical approaches and intuit the "feel" of good statistical inference, and (2) learn how to conduct basic empirical work in economics and become more comfortable with using data to tell a story. To understand proper *research design*, we'll explore common methods in modern econometrics through example. The course focuses on regression analysis, a tool that allows us to control for confounding factors embedded in real-world data and isolate causal effects. I aim to provide an exciting take on an interesting field while supplying you with the tools to become successful analysts.

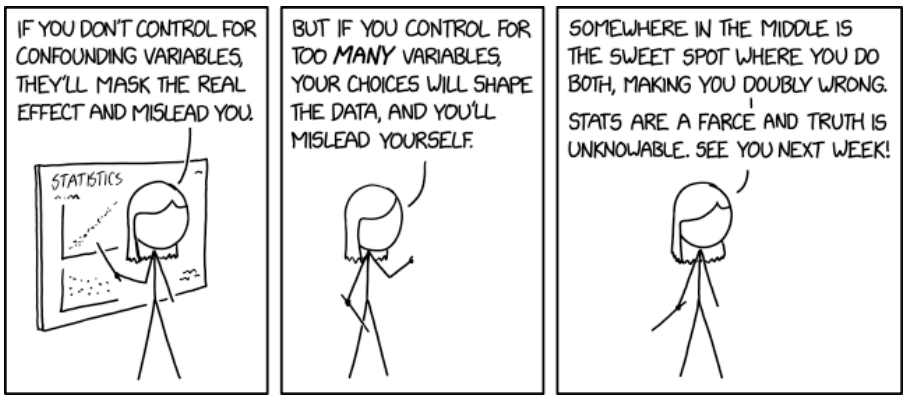

Figure 1: *Confounding Variables*, XKCD (12/27/2021).

[†]As the semester goes on, I may change the contents of this syllabus regarding the schedule, grading, or other details.

Prerequisites and Expectations

The prerequisite courses are Intermediate Micro/Macro (Econ 251 and 252), Statistics (Math 105 or equivalent), and Calculus (Math 161 or higher). Success in this course will in no small part be due to your previous mastery of the concepts learned in these courses.

Another prerequisite is a healthy learning attitude. Don't expect to fully-understand the material upon your first [passive] exposure during lecture. ECON 375 is one of those classes that takes a lot of effort outside of the classroom. Learning is supposed to take time, and it is supposed to be uncomfortable! Struggling with a concept is a precursor to truly understanding something. If you acknowledge that this will happen from time to time, there's suddenly no reason to feel frustrated, and that makes it easier to enjoy the process of figuring something out. I hope to show you that *'Metrics* is both a challenging *and* incredibly fun subject to explore during your time at Colgate.

You should aim to become more independent over the course of the semester. For example, before coming to me or your tutors, reflect on what you have tried and write down what didn't work and *why you think it didn't work*. You'll either spot the issue yourself, or come to me with your problem clearly identified. This strategy will keep you from giving up at the first sign of struggle.

You will be learning how to conduct empirical analyses in `STATA`. With respect to coding, independence becomes invaluable when debugging `STATA` problems because your bugs have been squashed many times before (and answered online in great detail). If you put in the effort into learning how to describe your coding problems, it'll become much easier to find those solutions.[†]

You're responsible for wanting to become (and becoming!) a competent econometrician. I can't accomplish this for you. Consider my role in the class. I am like your personal trainer: I provide you with your workout regimen (the lecture content, assignments, readings, and direction), and you choose how you want to follow that plan.[††] If you play an active role in class, regularly review your notes outside of class, and keep up with office hours, you're setting yourself up for success.

Textbook(s)

Most of our instruction will be inspired by the following book:

Mastering Metrics: The Path From Cause to Effect, by Joshua D. Angrist and Jörn-Steffen Pischke (2014)

The book is required for the course. You will be expected to have read the appropriate sections *before* the relevant lectures. The best way to follow along in the class is to have read the readings in advance. From time to time I may also assign other readings for certain lectures or homework assignments. For the reading schedule, see the course outline.

[†]You'll soon discover for yourself which resources you learn best from. `Google` is very useful here. As are the official `STATA` guides. `Wikipedia` is a great starting point for reviewing certain theoretical concepts. `StackExchange` provides invaluable discussion on some of the harder ideas. `Youtube` has hundreds of lessons and tutorials that could augment my lectures. On top of all of this, your classmates and tutors are also great coding resources.

[††]I'll be using `Moodle` to upload any resources (notes, readings, assignments, grades, etc.) we will be using throughout the course, and you'll submit assignments via `Gradescope`.

There is another [free!] book that I'd like to bring to your attention:

Causal Inference: The Mixtape, by Scott Cunningham (2021)
Link: `mixtape.scunning.com`

I will occasionally point to parts of *The Mixtape* that might complement the content in my lectures or in *Mastering 'Metrics*. In general, I think this newer book is a little tougher than *Mastering 'Metrics* on the math side of things but it explains some econometric concepts incredibly well and in very straightforward language. It also dives deeper into each topic, and with more examples than I could ever cover in class, so this book might be a huge boon for you.

Labs (and prelabs)

There will be eight labs on select Fridays throughout the semester (The other Fridays are for sleeping in, although the lab space will be reserved for you every Friday should you want to take advantage of this resource for your research paper or other projects). Lab sessions are designed to give practical experience with `STATA` and complement my [more theoretical] lectures. I will not be spending much time on `STATA` during class—it will be up to you to learn how to use this statistical software during these lab sessions (and office hours). I recommend working on these assignments in groups of two or three, although you must submit your own lab reports and code.

Your submissions will be uploaded to `Gradescope` as a single PDF file. I want to see a clean and professional-looking walkthrough answering each of the questions in a lab, using *curated* regression output and neat tables/figures (i.e. not gobs of copy-and-pasted lines from the `STATA` log). Your tutors shouldn't struggle to understand your writing or check your analysis/code. As an appendix, you should submit code for a single `STATA` ".do" file that would produce answers to all of the lab questions if I was to copy and paste into the `STATA` editor and click "execute." This code should be well-documented (headings, ownership, comments, etc.) so that anyone with some `STATA` experience could read and understand what it does (you should want to do this anyway, because you can usually reuse code from one week to the next).

Don't expect to complete each lab in the first two hours; coding-wise, this is likely, however, your full analysis will require a bit more effort (consider that this effort doesn't have to wait until Friday!). Labs will be due the following Tuesday before class, and will be submitted to `Gradescope` by 8:30am. Late labs will be accepted for two additional days for half credit.

You will complete a "pre-lab" before coming to each lab session. For those not familiar with the concept, pre-labs are shorter homework assignments that ensure you can hit the ground running during the designated lab time and make the best use of the resources available (the lab computers, and your tutors, in particular). You'll submit prelabs to `Gradescope` by 8am on lab days, and these may not be submitted late. Like the labs, each student must make their own submission.

Your *research paper* is another sort of lab report, although a more involved one—as it is on a topic of your choosing, relies on your own empirical direction, and is much more open-ended. We will build up to your final product together over the second half of the semester.

Lab Tutors

Tutor	Email	Office Hours (Persson 209)
Connor Appleyard	cappleyard@colgate.edu	Sunday, 6-7pm
Kaitlin Connelly	kconnelly@colgate.edu	TBD

I think Connor and Kaitlin are going to make great tutors, and I expect you to seek their help if you need it. They are primarily available to assist you with `STATA` and the labs, although they can help you with other parts of the course material as well.

Please remember that your tutors are also students—and that they have to balance this extra responsibility with their other course demands. Don't expect them to always be available to answer your questions, even if they happen to be awake when you happen to be asking for help. The best times to ask them questions are during labs and office hours. While emailing is acceptable, try to be aware of how often you expect something of them via this mode of communication (and allow them plenty of time to respond). Lastly, please don't ask anything of Connor or Kaitlin via informal means like texting or TikTok or walking over to their houses. If you happen to know them outside of ECON 375, be mindful of how those relationships interplay with your formal roles in this class.

Examinations

I design exams to evaluate how well you keep up with the assigned material and re-emphasize important points. They are opportunities for both feedback and additional learning. I hope to make you think critically about problems that you haven't faced before. That being said, you will find that the [three] exams are highly correlated in content and style with the class notes and labs/prelabs. The course is naturally cumulative, but each exam covers around a third of the material.

In place of a final exam, there will be a *lab practical*. This is an open-book, open-notes, open lab-work, and timed lab assignment, that requires you to demonstrate your ability to *produce* econometric analysis, and not just be a mere consumer. This will be in the form of a walk-through like the labs and will validate your learning of the `STATA` programming language over the semester.

Grading

I don't give grades, you earn them. Further, I don't judge your performance relative to your peers (i.e. curve your grades) during the term in order for you to have the clearest signal about your performance. I look for a proven understanding of the material via the following:

12%	Prelabs (1.5% each, x8)
12%	Labs (1.5% each, x8)
45%	Exams (15% each, x3)
15%	Research Paper
11%	Lab Practical

Conduct

These are just a few more things to keep in mind that I put in all my syllabi:

- Missing lecture isn't the end of the world. If you feel sick, please do not come to class. I don't need an excuse, but please be in touch if you will be out for multiple days.
- When asking for help outside of class, please be able to show how you have approached your problem. Simply asking for an answer is not a productive use of our time. I hope to facilitate critical thinking, and that takes effort on everyone's part.
- While I'll be accessible by email, I strongly prefer communicating during class/office hours. Regarding boundaries, I do not plan on answering emails late at night or on the weekend.
- I will not tolerate academic dishonesty. Colgate University's Academic Honor Code (`here`) requires instructors to report any suspected cheating, plagiarism, or other misconduct.
- You do not have permission to publish my course materials (online or otherwise). I don't want to see my work hosted somewhere like `CourseHero` (see the Academic Honor Code above).
- You are responsible for your technology problems. Be sure to submit assignments well ahead of the due date if you want to be sure that your submission is received/the proper format/etc.
- Please be respectful to your classmates. Refrain from talking during class if it is not relevant to lecture or discussion. Cell phone or tablet use should not detract from your ability to follow along with class. No activity on your part should undermine the efforts of other students.
- I can't recommend Colgate's Writing and Speaking Center (`link`) enough if you want to work on clear and coherent communication. Focused writing takes practice, and college is a great time to put in the hours.
- If you have any problems with this course or any other matters that may affect your work in this course, or you simply need someone to talk to, please contact me sooner rather than later. In cases where I'm not the appropriate resource, please seek support from Colgate's Counseling and Psychological Services (`link`). Reaching out to them is never a bad idea.
- If you have a learning disability or a physical disability that requires accommodation, please let me know as soon as possible. For more information on contacting the Office of Disability services, check out the Colgate University Disability Disclosure Statement that I have uploaded to `Moodle` under the syllabus tab.
- Colgate University is a diverse community of individuals with many perspectives and identities. In order to create an inclusive and intellectually vibrant community, we must understand individual differences and common ground. Colgate University's report on Academic Freedom and Freedom of Expression (`link`) reflects the ideals I seek to uphold in this class.

Course Outline

Linear Regression and Identifying Causality

Background Probability and [Frequentist] Statistics ..... Tuesday, 1/25

Mastering 'Metrics: Intro

Concepts: random variable, sampling, data-generating process, probability distribution, parameter, p-values, confidence intervals, inference

More on Statistical Inference ..... Thursday, 1/27

Causal Inference Mixtape: Probability and Regression Review (Skim)

Concepts: sampling variance, estimators, law of large numbers, consistency, unbiasedness, the central limit theorem, t-statistics, hypothesis testing and statistical significance

Notation, Notation, Notation ..... Tuesday, 2/1

Causal Inference Mixtape: Probability and Regression Review (Skim)

Concepts: data reduction, summation operator, properties of expectation/variance/covariance

The Potential Outcomes Framework ..... Thursday, 2/3

Mastering 'Metrics: Chapter 1, Chapter 1 Appendix

Reading: NYT analysis of the Nike Vaporfly (link: NYT)

Concepts: treatment and control groups, dummy variables, counterfactuals, selection bias

Lab #1: Introduction to `STATA` ..... Friday, 2/4

Placeholder day (potential travel) ..... Tuesday, 2/8

Randomized Controlled Trials ..... Thursday, 2/10

Mastering 'Metrics: Chapter 1, Chapter 1 Appendix

Concepts: random assignment, conditional expectation, average treatment effects

Lab #2: Randomized Controlled Trials ..... Friday, 2/11

Regression Anatomy ..... Tuesday, 2/15

Mastering 'Metrics: Chapter 2 Appendix

Concepts: curve-fitting, residuals, optimization, the least-squares estimator

Regression Physiology ..... Thursday, 2/17

Mastering 'Metrics: Chapter 2 Appendix

Concepts: Gauss-Markov assumptions, distribution of regression parameters

Linear Regression and Identifying Causality, cont'd.

Exam 1: Statistical Inference and Randomization ..... Tuesday, 2/22

Directed Acyclic Graphs ..... Thursday, 2/24

Causal Inference Mixtape: Directed Acyclic Graphs
Concepts: counfounders, colliders, backdoor paths, conditioning

Lab #3: The Mechanics of Regression ..... Friday, 2/25

Reading: Speeding, Coordination, and the 55 MPH Limit (Lave, 1985)

Causal Thinking and Regression ..... Tuesday, 3/1

Mastering 'Metrics: Chapter 2
Concepts: omitted variables bias, control variables, identifying variation

Regression as a Matchmaker ..... Thursday, 3/3

Mastering 'Metrics: Chapter 2
Concepts: selection on observables, omitted variables bias, causal vs control variables

Lab #4: Regressions for Causality ..... Friday, 3/4

Panel Data and Fixed-Effects ..... Tuesday, 3/8

Mastering 'Metrics: Chapter 2, Chapter 2 Appendix
Concepts: repeated observations, longitudinal and panel data, unobserved heterogeneity, pooled OLS, fixed-effect regrssion, identifying variation

Exam 2: Regression and Causality ..... Thursday, 3/10

(Not actually a lab but a paper topic brainstorming session) ..... Friday, 3/11

Spring Break Week ..... Tuesday, 3/15

Research Design

Instrumental Variables ..... Tuesday, 3/22

Mastering 'Metrics: Chapter 3

Concepts: "partialing out" procedure, instrument, proxy variable, two-stage least squares, exclusion and independence of an instrument

IV and Validity ..... Thursday, 3/24

Mastering 'Metrics: Chapter 3

Concepts: intent-to-treat effect, local average treatment effect, internal and external validity

(Not actually a lab but a research paper help session) ..... Friday, 3/25

Regression Discontinuity: Design ..... Tuesday, 3/29

Mastering 'Metrics: Chapter 4

Concepts: running variable, threshold, jump, "local randomization," natural experiments

R_2D Two: Electric Boogaloo ..... Thursday, 3/31

Mastering 'Metrics: Chapter 4

Concepts: bandwidth, local linear regression, polynomials and splines

Lab #5: Regression Discontinuity ..... Friday, 4/1

Fuzzy Regression Discontinuity ..... Tuesday, 4/5

Mastering 'Metrics: Chapter 4

Concepts: "ocular" econometrics, imperfect treatment, instrumental variables again

Differences-in-Differences: Design ..... Thursday, 4/7

Mastering 'Metrics: Chapter 5

Concepts: panel data, common (parallel) trends, time trends, interaction variables

(Not actually a lab but another research paper help session) ..... Friday, 4/8

Deeper into "Diff-in-Diff" ..... Tuesday, 4/12

Mastering 'Metrics: Chapter 5

Concepts: placebo variables, serial correlation, clustered errors

DiD: Examples ..... Thursday, 4/14

Reading: Racial Discrimination among NBA Referees (Price and Wolfers, 2010)

Reading: Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education (Bursztyn, Egorov, and Janzen, 2017)

Lab #6: Differences-in-Differences ..... Friday, 4/15

Beyond Causality

Predicting the Future ..... Tuesday, 4/19

Concepts: forecasting, longitudinal data, autocorrelation, trends and seasonality

Single-equation Time Series Econometrics ..... Thursday, 4/21

Concepts: autoregressive/distributed lag models, Granger causality, Newey-West estimator, unit roots, random walks, cointegration

(Not actually a lab but one last research paper help session) ..... Friday, 4/22

Exam 3: Research Design ..... Tuesday, 4/26

No class (at Virginia Tech) ..... Thursday, 4/28

Lab #7: Time Series Econometrics ..... Friday, 4/29

Discrete Choice ..... Tuesday, 5/3

Concepts: binary outcomes, logistic function/regression, odds, predicting probabilities

Maximum Likelihood Estimation ..... Thursday, 5/5

Concepts: likelihood function, extrema

Lab #8: Logistic Regression and Binary Outcomes ..... Friday, 5/6

Lab Practical, Section B (9-11am) ..... Monday, 5/9

Lab Practical, Section A (9-11am) ..... Tuesday, 5/10