

ECO 439
Methods for Empirical Microeconomics

University of Toronto

Department of Economics
Winter 2019

Course Description

This course is directed at undergraduate students anticipating that they may attend graduate school in economics. This pertains especially (but not exclusively) to students expecting to work or take graduate classes in labour, development, and public economics. It is a course in empirical modeling and applied econometrics. The tools covered in the course, however are central to those used in empirical labour economics, as well as other applied microeconomics fields like development and public economics. The focus will be on the identification of casual relationships using regression-based analysis. Empirical examples will be drawn from recent work in labour, development, and public economics.

Instructor

Instructor: Arthur Blouin
Email: a.blouin@utoronto.ca
Office: 150 St. George Street, #305

Office Hours: Monday, 10:00 to 11:00

Pre-requisites

Applied Econometrics (ECO375/374)

OR

80% in:
Microeconomic Theory (ECO200/204/206)

AND

Quantitative Methods in Economics (ECO220/227)

Meetings

Lectures are Tuesdays, 3:10 to 5:00, in TF202. We will NOT make use of the Thursday time slot unless something unexpected happens and we need to make-up time. I will announce ahead of time whether we will be using the Thursday slot in a particular week, so if it is not announced, there is no Thursday class.

Readings

The core lecture material is based on:

Joshua D. Angrist & Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

This can be purchased from various online booksellers.

In addition to the textbook, a central part of the course will be selected journal articles that illustrate the various empirical strategies and methods that we will be discussing. The articles will be drawn broadly from empirical microeconomic fields, and the course will therefore have “economic content” in addition to the focus on applied econometrics. A more complete list of the readings is listed below.

Website

The course website (on Quercus).

I will post the slides from lectures on Quercus. I tend to use Quercus extensively as a means of communication with the class, so I recommend you check the announcements regularly.

Email Policy

Please feel free to email me questions or comments pertaining to the course, with the following proviso:

The answer requires a one or two-line response (maximum). It is my experience that email is an inefficient way to discuss economics. Questions that require more than one or two-line answers are more appropriate for office hours.

I will normally reply to emails within 24 hours, except on weekends.

Evaluation

- Term paper (30%)
- (4) Problem sets (20% - 5% each)
- Midterm exam (20%)
- Final exam (30%)

Papers and problem sets are penalized by 10% per day late.

Term Paper:

The term paper will test a hypothesis. It doesn't have to be strictly economics; I think of this as a stats class applied to economics. In the past I've received a lot of papers testing sports hypotheses, for example. You are expected to clearly outline some hypothesis, develop an empirical strategy and implement your strategy to test your hypothesis. You will be graded on the methods used and your ability to understand how

close you have come to a causal estimate. The actual estimates you get don't matter for your grade - i.e. I don't care if nothing is statistically significant and you can't make any conclusions from your work – just interpret things properly.

I won't be strict about length of the paper, but there is no need for it to be 20+ pages, and you'll likely have a tough time doing what you need to do in 4 or fewer pages. Look at papers published in the *American Economic Review: Paper and Proceedings* as a guide for format and length.

The following is a guideline of what I expect:

- 8-12 pages of text plus 3-4 tables of supporting evidence and a page citing sources.
 - Tables include a Summary Statistics table, possibly some sort of balancing table, and 2-3 tables with regression results.
- An introduction that outlines your hypothesis and basic argument
- A (short) data/sources section describing where your supporting evidence is coming from
- A results section where you describe how your evidence relates to your hypothesis.
 - In this section it is crucial to document and discuss both the strengths AND weaknesses in the evidence you present. There are no perfect papers, even seminal papers in the field rarely (if ever) perfectly clinch arguments. Good papers outline weaknesses and discuss how concerned we should be with these weaknesses when making inferences.

The term paper is due at 5:00pm on Tuesday April 2nd, 2019.

- Hard copy submissions only.
- You can either hand it in during/after class or slide it under my door or hand it in (and get it time stamped) at the front desk of the economics department (Max Gluskin House).
- Emailed copies of the paper will be ignored, and papers that are not time stamped and not in my hands by 5:00 (or shortly thereafter if you slide it under my office door during class) will be considered late.

Problem sets:

There will be 4 problem sets applying techniques learned in class to data. For this you will need Stata, the small version will be fine. You'll have to run some analysis in Stata and interpret your findings in the .do file with comment tags e.g.: `/*COMMENT*/`

If you aren't comfortable with Stata, please let me know at the beginning of the year. There is a Stata TA for the course, and if there is sufficient demand we can do one or two Stata tutorials in the Thursday tutorial hour.

I will post a problem set on Quercus and announce in class when I do. Problem sets will be posted as soon as we've covered enough material for you to be able to do them. You'll have a week from when I announce that the problem set is posted to complete it. Each one is 3-4 questions and should take you about one to two hours to do.

Please email completed problem sets to the following email address: ec439@gmail.com (NOTE: it's not a typo, it's "ec" and not "eco," I get asked about this every year: I set this

email address up years ago when I joined U of T, and the school I used to be at used EC as a course naming convention, and I was still defaulting to that)

There are workstations with Stata on them in the Rotman computer lab, otherwise you can get a 6-month student version of Stata online for \$45 here:
<http://www.stata.com/order/new/edu/gradplans/student-pricing/>

Mid-Term:

The mid-term is scheduled for February 12th at 3pm-5pm in class. We will aim to start as soon as possible after 3pm (this depends on if we're in a room with a class before us) and ideally we will start right at 3pm, and not 3:10. The exam is 2 full hours, so for example if we don't get started until 3:10, you will have until 5:10.

The mid-term will cover all material up to and including the lecture on February 5th.

Missing mid-term or assignment deadlines:

A grade of 0 will be given to students who do not write the midterm exam unless an appropriate and convincing note is received within one week of the missed test, explaining why the test was missed.

All assignments will be penalized at 10% per day late.

If the student misses a test due to illness:

- The note must be provided using the University of Toronto medical certificate. No other documentation will be accepted. You can find a copy of the form here: <http://www.healthservice.utoronto.ca/pdfs/medcert.htm>
- The form must be completed by a Medical Doctor, and include the doctor's OHIP registration number.
- Only original notes will be accepted. I will not accept photocopies or emailed certificates.
- The note must clearly state that on the date of the test, the student was too sick to write the test.
 - Notes that simply state that a doctor saw the student on the date of the midterm will not be accepted.
 - Illness before the test is not sufficient grounds for missing the test.
 - Notes that state that a Medical doctor saw the student in the days following the test will not be accepted.
 - Nor will I accept notes that indicate that the student would have performed "sub-optimally".
- **To comply with these requirements, it is expected that the student will have met with the doctor on the date of the test.**
- **The student must email me the day of the test to indicate that they will not be able to write the test.**
- I will review each sick note to determine whether there are sufficient grounds for a student to be excused from a test. Part of this review process may include meeting with the student, and/or following up with a physician.

If the student misses a term test due to another excused absence (e.g., funeral, car accident):

- The note must be accompanied by a note from a responsible third party that I can verify *in order for excusal to even be considered*. The validity of the excuse will also be evaluated by academic staff in the Department of Economics.

If a student has been excused from the midterm exam, he or she will be permitted to write a make-up test. The make-up test will be held a week or two after the midterm (tbd). Consistent with university policy, there is no make-up test for the make-up test. No medical excuses or scheduling conflicts will be accepted, and a grade of zero will be applied if a student fails to write the make-up test.

Accessibility Policy:

If you need an extension for an assignment or miss a test for accessibility reasons, the following policy applies:

- A current form documenting registration with accessibility must be shown to me.
- The student's accessibility adviser must contact me clarifying that the student requires an extension for an accessibility related reason.
 - Please note: it is not enough for the accessibility advisor to state that the student continues to be registered with accessibility, I need confirmation from the accessibility advisor that the missed test or assignment deadline is due directly to an accessibility related reason.
- I do not need to know the reason for registration with accessibility, nor is it appropriate for a student to provide these details in an effort to persuade me to grant an extension. It's not that I don't care, it's that I'm not a medical doctor and am not qualified to assess medical or mental health. This is confidential information between the student and their medical/mental health support.
- In general, I will follow the recommendation of the accessibility advisor on the length of an appropriate extension, provided they are able to claim that the missed test or assignment deadline was directly due to whatever the reason for registration with accessibility was in the first place.

Extensions:

I will give extensions for large, exogenous negative productivity shocks, but the following reasons are typically not sufficient grounds for an extension for any assignment:

- work would be improved if the student had longer to do it.
- student has lots of other assignments due around the same time, or is in any other way having difficulty managing time.
- Student is having a difficult time with the assignment.

Unless something unexpected happens I'm typically not generous with extensions. If anyone anticipates having many papers due in the last week of class (as is the case in this class) they do have the option of completing the paper at any other time during the

term. I think that time management is an important skill, and a failure to manage deadlines is not sufficient grounds for an extension.

Final Exam:

The final will be governed by the University's rules for missing final exams.

Planned Coverage

We will follow the material outlined in Angrist & Pischke very closely, with some recent innovations:

1. Introduction to the "Experimental Ideal" (Chapters 1 and 2);
2. Detailed review of Ordinary Least Squares and Regression analysis (Chapter 3);
3. Matching (Chapter 3.3);
4. Instrumental Variables (Chapter 4);
5. Regression Discontinuity and Regression Kink Designs (Chapter 6);
6. Panel Data and Differences-in-Differences (Chapter 5);
7. Machine Learning
8. Issues with Standard Errors (Chapter 8).

Preliminary List of Readings

The following is a list of the key parts of the text, and associated journal articles that we will be (mostly) covering in class (or that are discussed in some detail in Angrist and Pischke). The articles that form the basis of the assignment will be ADDED to this list (and they are required readings for the entire class).

In addition to the presentation in Angrist and Pischke, a denser, but clear and comprehensive discussion of the course material is provided by:

Guido Imbens and Jeffrey Wooldridge (2009): "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47:1, pages 5-86.

1. Introductory Material

Angrist and Pischke, Chapters 1 and 2

2. The Regression Model

Angrist and Pischke, Chapter 3, Sections 3.1 and 3.2

3. Matching

Angrist and Pischke, Chapter 3, Section 3.3

LaLonde, Robert (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review* 76, September, pp. 604-620.

Ashenfelter, Orley (1978): "Estimating the Effect of Training Programs on Earnings," *The Review of Economics and Statistics* 60, pp. 47-57.

Ashenfelter, Orley, and David Card (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs on Earnings," *The Review of Economics and Statistics* 67, pp. 648-66.

Dehejia, Rajeev, and Sadek Wahba (1999): "Causal Effects in Nonexperimental Studies: Re-evaluating the Evaluation of Training Programs," *JASA* 94.

Smith, Jeffrey, and Petra Todd (2001): "Reconciling Conflicting Evidence on the Performance of Propensity Score Matching Methods," *American Economic Review* 91, May.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica* 71:4, 1161-1189.

Imbens, Guido W. (2000). "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika*, 87, 706-710.

Lechner, Michael (2002a). "Programme Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labour Market Policies." *Review of Economics and Statistics*, 84, 205-220.

Lechner, Michael (2002b). "Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods." *Journal of the Royal Statistical Society, Series A*, 165, 59-82.

Abadie, Alberto, and Javier Gardeazabal (2003). "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1), 113-32.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490), 493-505.

4. Instrumental Variables

Angrist and Pischke, Chapter 4

Angrist, Joshua (1990): "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review*.

Angrist, Joshua, and Alan Krueger (1991): "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106.

Imbens, Guido, and Joshua Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, Vol. 62, No. 2, pp. 467-475.

Angrist, Joshua (1998): "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*.

Bound, John, David Jaeger, and Regina Baker (1995): "Problems with Instrumental Variables when the Correlation Between the Instruments and Endogenous Variable is Weak," *Journal of the American Statistical Association*.

Card, David (1999): "The Causal Effect of Education on Earnings," Chapter 30 in Ashenfelter, Orley, and David Card (eds.) *Handbook of Labor Economics*, Volume 3.

Oettinger, Gerald (1999): "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors," *Journal of Political Economy*, 107(2).

Deaton, Angus (2010): "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, 48, pages 424-455.

Imbens, Guido (2010): "Better LATE than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, 48, pages 399-423.

5. Regression Discontinuity, Regression Kink Designs, and Bunching

Angrist and Pischke, Chapter 6

RDD

Lee, David, and Thomas Lemieux (2010): "Regression Discontinuity Designs In Economics," *Journal of Economic Literature*, 48, pages 281-355.

Angrist, Joshua, and Victor Lavy (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, pp. 533-575.

Lee, David (2008): "Randomized experiments from non-random selection in U.S. House elections," *Journal of Econometrics*.

Lemieux, Thomas, and Kevin Milligan (2008): "Incentive effects of social assistance: A regression discontinuity approach," *Journal of Econometrics*.

Imbens, Guido W., and Karthik Kalyanaraman (2012). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*, 79(3), 933-959.

RKD

Guryan, Jonathan (2001). "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts." NBER Working Paper 8269.

Dahlberg, Matz, Eva Mork, Jorn Rattso, and Hanna Agren (2008). "Using a Discontinuous Grant Rule to Identify the Effect of Grants on Local Taxes and Spending," *Journal of Public Economics*, 92(12), 2320-2335.

Card, David, David Lee, Zhuan Pei, and Andrea Weber (2012). "Nonlinear Policy Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design." NBER Working Paper 18564.

Bunching

Saez, Emmanuel (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2, 180-212.

Kleven, Henrik J., and Mazhar Waseem (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 128, 669-723.

6. Panel Data and Differences-in-Differences

Angrist and Pischke, Chapter 5

Card, David (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 1990.

Ashenfelter, Orley, and Alan B. Krueger (1994): "Estimates of the economic returns to schooling from a new sample of twins," *American Economic Review* 84, (5) (December 1994): 1157-73.

Duflo, Esther (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), 795-813.

7. Machine Learning

Mullainathan, Sendhil, Jann Spiess (2017). "Machine Learning: An Applied Econometric Approach" *Journal of Economic Perspectives*, 31(2), 87-106.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, Sendhil Mullainathan (forthcoming). "Human Decisions and Machine Predictions." *Quarterly Journal of Economics*, qjx032.

Varian, Hal (2014). "Big Data: New Tricks for Econometrics" *Journal of Economic Perspectives*. 28(2), pp. 3-28.

8. Issues with Standard Errors

Angrist and Pischke, Chapter 8

Moulton, Brent (1986): "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics* 32, pp. 385-97.

Bertrand, Marianne, Esther Duflo, Sendhil Mullainathan (2004). "How Much Should We Trust Difference-in-Difference Estimates?" *Quarterly Journal of Economics*, 119(1), 249-75.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics*, 90, 414-427.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller (2011). "Robust Inference with Multi-Way Clustering," *Journal of Business and Economic Statistics*, 29(2), 238-249.

Cameron, A. Colin, and Douglas L. Miller (2015). "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50(2), 317-73.

Imbens, Guido W., and Michal Kolesar (2012). "Robust Standard Errors in Small Samples: Some Practical Advice." NBER Working Paper No. 18478.

Ibragimov, Rustam, and Ulrich K. Müller (2014). "Inference with Few Heterogeneous Clusters." *Review of Economics and Statistics* (forthcoming).