UNIVERSITY OF TORONTO

ECO372H1 Data Analysis and Applied Econometrics in Practice

Winter 2021

Sections L0101, L9101, L0201 & L9201— Instructor : Patrick Blanchenay

This version: **2021-01-12**. [Latest version]

Key Information

Course site: Quercus (all announcements and material)
Calendar: see Quercus homepage

Instructor: Patrick Blanchenay
Head TA: Marc-Antoine Chatelain

Drop-in Office hours: Tuesdays 9.00-10.00 on Bb Collaborate (access from Quercus)
Communication: see Communication below
Piazza: Access on left tab from Quercus

All announcements will be made *using* Quercus. All material will be posted on Quercus.

Description

ECO372 Data Analysis and Applied Econometrics in Practice is an intermediate level course in econometrics for students at the University of Toronto (St-George campus). Its goal is to equip students with a modern approach to data analysis and econometrics, focussing on the use of data to answer causal questions. Students 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. Students will learn about applications of these techniques in academic research. Students will also put these techniques in practice and gain familiarity with Stata, one of the most widespeard statistics software in economics.

Learning Outcomes

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

1. Understand the notion of causality, and its importance in applied empirical research.
2. Identify five strategies that can be used 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, typical use, and limitations, and know how to interpret their quantitative results.
4. Read unfamiliar research papers and understand the research question and basic estimation strategy / econometric model used to answer it.
5. Interpret and comment on tables of estimated coefficients from a wide range of econometric models, in various formats.
6. 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.
7. Articulate short well-crafted arguments to answer questions regarding each of these methods.
8. Reproduce key results of empirical papers using Stata.
9. Apply these methods to actual datasets, using Stata programming language.

Course delivery & scheduling

The course is fully online. Some activities take place during your scheduled classtime and you are expected to be available.
Our workweeks run from Monday to Friday (see Quercus homepage for details). Content is organized by workweek. Every Monday, a new module is made available. A module typically comprises pre-recorded (asynchronous) lecture videos, required readings, as well a tutorial handout.

For each week, you are expected to complete the readings and watch the lecture videos, work on the tutorial, participate in group discussions, and complete quizzes and assignments by their due dates. I expect a time commitment of 10 hours a week. Assignments are made available for several days and are not timed. Quizzes and the final assessment are timed but can be taken at any time during their respective window.

Watch the lectures and do the readings first. Then work on the tutorial. Pre-recorded (asynchronous) tutorial walkthrough videos are posted on the Wednesday.

Office hours will take place online each week, at a time different than the time slot (see details on Quercus). Hands-on help (office hours/Stata help) will be delivered online during our scheduled class timeslo. For sensitive matters, online meetings can be arranged.

The course is organized around 6 topics.

Week 1	Causality & Statistics
Weeks 2-3	Random Assignment (Randomized Controlled Trials)
Weeks 4-5-6	Linear Regressions using OLS
Weeks 7-8	Instrumental Variables (IV)
Weeks 9-10-11	Difference-in-differences and Panel Data
Week 12	Regression Discontinuity Design (RDD)

Academic Prerequesites

The course prerequisites and exclusions are listed here: https://fas.calendar.utoronto.ca/course/eco372h1 .
I cannot waive prerequisites. An administrator will remove anyone missing prerequisites.

Technology requirements

You must meet the minimum technology requirements for online learning required by the UofT listed at https://www.viceprovoststudents.utoronto.ca/covid-19/tech-requirements-online-learning/ . Please also read the recommended accessories and internet connectivity sections. For this course, you must have:

- Access to a laptop or desktop computer with a working microphone and webcam. A phone is not an acceptable substitute as some required components may not be accessible on smartphones.
- The laptop must be able to run Stata. (See the syllabus section on Stata.)
- Access to stable, high-speed internet and reliable electricity.
- The technology and knowledge to scan or photograph your handwritten work and convert it into small PDF, JPG or PNG files.

Sometimes things can go wrong with technology. Please be proactive—maintain regular backup copies of your files, use antivirus software, and submit your work well before deadlines in case of technical difficulties. Common issues like computer viruses, crashed hard drives, lost or corrupted files, incompatible file formats, faulty internet, etc. are not acceptable reasons for a deadline extension.

Required material

Textbook

The required textbook is *Mastering 'Metrics* by Angrist & Pischke (Princeton University Press, ISBN: 978-0-691-15284-4). It is cheap and available at UofT Bookstore, as well as on Amazon, Indigo and many other online platforms. There are copies available at UofT libraries, but even if you are in Toronto, I strongly recommend that you buy your own copy. Not only is it an engaging and accessible book, it is also one that can stay useful for a lifetime (even if you opt for a career outside academia).

The course will follow the themes of the textbook, but not in the original order.

You are free to supplement with other textbooks of your choice; I find Stock and Watson's *Introduction to Econometrics* (Pearson) excellent on Regression analysis, and Instrumental Variables, but it has little Difference-in-Differences and Regression discontinuity design.

Readings

With each week, I will post the required readings, which will include parts of the textbook, and additional readings (typically research papers) that I will upload to Quercus.
All assigned readings are examinable, and you should read the material in parallel to watching the lectures. Research papers are examinable in the sense that you are expected to understand and comment on the methodology they use. You are not expected to remember specific results.

Stata 14+

One objective of the course is to get students to perform analyses using statistical software. Some tutorials will provide hands-on experience to Stata, one of the most widely used statistics software in the world.

Stata is a proprietary commercial software, available on Windows, macOS and Linux. Stata comes in successive versions. You need version 14 or above. For each version, Stata comes in several "flavours" of increasing memory capacity, but for the purpose of this course, the simplest Stata/IC is sufficient. As of August 2020, a 6-month license to Stata/IC for students costs 48 USD (make sure to click on the 6-month tab). (You are free to buy a more expensive flavour or a longer license, but it is unnecessary for this course. Do NOT buy Small Stata.) If you alreay have Stata version 14 or above, you do not need to buy the latest version, as most commands we will use haven't changed. Stata 13 is unable to open datasets provided for the tutorials and assignments (see also FAQ: Can I use Stata 10/11/12/13?).

Stata is also available on computers of the Map and Data Library room at Robarts Library; it is your responsibility to adequately transfer your files. See also FAQ: Do I need to buy Stata? Can I use R instead? and Am I going to become a Stata/programming wizard?

Evaluation

Marking scheme

Evaluation	Due Date	Percentage of final grade
3 Group Discussions	Mon 22 Feb Fri 12 Mar Tue 30 Mar	Equally weighted, total 10%.
3 Quizzes	Thu 25 Feb Tue 16 Mar Thu 01 Apr	15% each, lowest of 3 is dropped. Total: 30%.
4 Assignments	Sun 31 Jan 6PM Sun 7 Mar 6PM Sun 21 Mar 6PM Fri 09 Apr 6PM	13% each, lowest of 4 is dropped. Total: 39%
Final Assessment	During 13-23 Apr 2021, exact date to be confirmed.	21%

Group Discussions (10%) and Quizzes (30%)

There are 3 Group discussions, each followed by a quiz. You are provided with an article to read, and some points to discuss.

Group Discussions: You will be assigned to a group of 4 students, and asked to record a 20-30min online discussion with your classmates. The video must be posted by the deadline. Your participation in that discussion will be individually graded based on the quality and relevance of your input. Not attending the group discussion with your classmates will earn a score of zero for that group discussion. Groups will vary from discussion to discussion. A late submission will result in a grade of zero for all participants in the group.

Quiz: Following each group discussion, you must individually take a quiz comprising short questions on the reading you discussed with the group. The quiz is timed (typically 80min), during a wider completion window. Collaboration is not permitted on the quizzes. You will be graded based on your understanding of the course concepts and how they apply to the assigned reading, as well as your clarity. The lowest of your 3 quiz scores is dropped. A late submission will result in a penalty of 20 points per day of lateness.

Assignments (39%)

There are four Stata assignments to hand in. These exercises involve a mix of algebra, questions based on data analysis using Stata, and writing about your analysis in a clear rigorous way.

For every assignment, you will be asked to submit three documents:

- a PDF containing the answers to your questions, with the adequate explanations and interpretations
- the Stata code (do-file) that you used for the analysis, duly commented;
- the log file automatically generated by your code.

You will have to upload these documents before the deadline to Quercus. Assignments that are submitted late will suffer a penalty of 10 points per day of lateness, starting immediately at the deadline (by the deadline). Files are automatically processed by Turnitin, a plagiarism detection software. See sections Academic Integrity and Turnitin.com.

You will be graded on the quality of your answers, with a big emphasis on clarity. The lowest of your 4 assignment scores is dropped.

Final Assessment (21%)

A final assessment will take place during the Final Assessment period (13-23 Apr); the final assessment will comprise multiple questions that are longer in nature to those of the quizzes, and can assess your in-depth understanding of the course material. The Final Assessment is timed during its completion window; you might be expected to upload answers to Crowdmark (detailed instructions will be provided). Stata is not used for the final assessment. Late submissions will result in a grade of zero.

Missed term work and Late Penalties

You are expected to complete all the work scheduled in the Evaluation section. Assessments dates and instructions are posted in advance, and it is your responsibility to ensure adequate time to complete the work and deal with any issues, including technical issues. This is particularly important for this online course.

Failure to submit an assessment will result in a grade of zero. Late submission may result in penalties or a grade of zero. Make sure to allow ample time for submission before the deadline; excuses such as: “the website is slow”, “I lost my Internet connection”, “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.

By dropping the lowest quiz, and the lowest assignment. The marking scheme already acknowledges the various (and hard to document) circumstances that may cause a student to write in bad conditions. No documentation is needed.

Accommodations for missing more work than already addressed in the previous paragraphs (missing more than one quiz or assignment) are extremely limited: (A) an ongoing and substantial injury, illness, or personal/family problem seriously affecting the student's ability to complete term work across all courses over an extended period of time, where the student's College Registrar writes to each professor after reviewing the documentation and meeting with the student; or (B) more than one conflict not related to injury, illness or personal/family problems where we are contacted by the student very far in advance (e.g. an athlete who notifies me in January of international competitions running of the periods of several assessments). In these limited situations, I will consider whether accommodations can still meet all course requirements or whether the student must be advised to drop the course and retake it when able to complete the required work.

For ongoing injury, illness, or personal issues, you are strongly encouraged to contact your College Registrar office for guidance and support, as soon as possible. Your registrar can then get in touch with me as appropriate.

Academic Skills Resources

Even the most seasoned, organized and dedicated student can benefit from speaking with a Learning Strategist to explore setting achievable goals, preparing for tests and time management strategies. You can schedule an appointment with a Learning Strategist at any time of the year.
https://sidneysmithcommons.artsci.utoronto.ca/can-a-learning-strategist-help-me/

Academic Integrity

All students, faculty and staff are expected to follow the University's guidelines and policies on academic integrity. For students, this means following the standards of academic honesty when writing assignments, citing and using source material appropriately, collaborating (or not collaborating) with fellow students, and writing tests. Ensure that the work you submit for grading represents your own honest efforts.

Unfortunately there has been cases of academic dishonesty in ECO372 in the past, and the students have received heavy sanctions, and all regretted. Speak to me or your TA for advice on anything that you are unsure about anything

Potential offences include, but are not limited to:

- Using someone else's ideas or words without appropriate acknowledgement. This includes verbatim copying of any lecture notes distributed by the instructor.
- Submitting your own work in more than one course without the permission of the instructor.
- Making up sources or facts.
- Misrepresenting your identity.
- Obtaining or providing unauthorized assistance on any assignment, including from your classmates.

I do encourage you to pay close attention to these sections on Perils and Pitfalls
http://academicintegrity.utoronto.ca/perils-and-pitfalls and Smart Strategies
http://academicintegrity.utoronto.ca/smart-strategies

Also, see the U of T writing support website at http://www.utoronto.ca/writing. Consult the Code of Behaviour on Academic Matters (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm) for a complete outline of the Universitys policy and expectations.

Turnitin.com

Normally, students will be required to submit their course assignments to Turnitin.com for review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the Turnitin.com reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of the Turnitin.com service are described on the Turnitin.com web site.

Communication

I do not use Quercus messaging tool; **please do not send any message via Quercus**. Instead we have two channels of communication: Piazza, and email.

Piazza

We use Piazza (https://q.utoronto.ca/courses/204242/external_tools/5714) to facilitate communication.
The TAs and I periodically check Piazza to ensure proper usage, flag some postings, and possibly answer some questions. However, Piazza's emphasis is on student-to-student Q&A. Questions and answers can be made anonymous to other students (but not to the teaching team).

Piazza is a complement to face-to-face interactions in office hours, class, TA tutorials, and study groups. Piazza has several advantages over email; first, it promotes student engagement by encouraging you to answer other students' questions, an excellent way of testing your understanding of the material; second, it allows questions/answers to be shared to all students, who can benefit from this positive externality.

Email

Email is not an appropriate forum for discussing course material details, which are better addressed on Piazza or through office hours. That said, email can be helpful on occasion, and within limits and I will try to reply to email provided your question(s) can be answered with a one or two sentence answer. I have more time to reply to email on Mondays and Tuesdays; if anything is urgent, you can also come to my office hour.

Answering email takes up valuable time that could be used more productively on improving the course; I will not answer about information that can be found in the syllabus or on the Quercus page (for example: test dates) nor questions about grades, nor questions about the course material (which should be posted on Piazza). Please do not send attachments and do not submit term work by email. **Please always include [ECO372] in the title of *any* email. Emails that do not include this will be ignored.**

The instructor's email address is: patrick.blanchenay@utoronto.ca

Email checklist:

- ❏ Your email is sent from your University address.
- ❏ Title includes "[ECO372] " followed by the subject of the email.
- ❏ The message is straight to the point and no-longer than 10 sentences.
- ❏ Your email does not contain attachment, unless I had specifically asked you to send something.
- ❏ Your signature includes your Student number, and both the name you are usually addressed by, and your ACORN name if it differs from the name you go by.

Student Well-Being and Academic Accomodations

Well-Being

University of Toronto aims at giving you an enriching learning experience, and has a number of resources to help you stay healthy and be well: http://studentlife.utoronto.ca/bewell These resources can be useful even if you are following the course remotely.

Sometimes things do not go as planned. **In case of emergency, call your emergency services (911 in Canada).** For ongoing injury, illness, or personal/family problems, or if you feel you are falling behind in your courses, you must contact your College Registrar immediately, who can then get in touch with your instructors. The earlier you do, the easier it is to find solutions.

There are also a number of resources in case you are feeling distressed, and many of these are accessible even if you're outside Canada: http://studentlife.utoronto.ca/feeling-distressed
Once again, the earlier you reach out, the easier it is to remedy the situation and find solutions. Do not wait until the end of the academic year.

Ongoing Learning Disability or Accommodation Requirement

Students with diverse learning styles and needs are welcome in this course. If you have an ongoing disability issue or accommodation need, you should register with Accessibility Services (AS) (http://accessibility.utoronto.ca) **at the beginning of the academic term**. (Without registration, you will not be able to verify your situation with your instructors, and instructors will not be advised about your accommodation needs.) AS will then assess your medical situation, develop an accommodation plan with you, and support you in requesting accommodation for your course work. Remember that the process of accommodation is private: AS will not share details of your condition with any instructor, and your instructors will not reveal that you are registered with AS.

For more information on services and resources available to instructors and students, please contact Tanya Lewis, Director, Director of Academic Success and Accessibility Services, at (416) 978-6268; tanya.lewis@utoronto.ca.
Accessibility services: http://studentlife.utoronto.ca/as/

Accommodations for Religious Observances

As a student at the University of Toronto, you are part of a diverse community that welcomes and includes students and faculty from a wide range of backgrounds, cultural traditions, and spiritual beliefs. For my part, I will make every reasonable effort to avoid scheduling tests, examinations, or other compulsory activities on religious holy days not captured by statutory holidays. Further to University Policy, if you anticipate being absent from class or missing a major course activity (like a test) due to a religious observance, please let me know as early in the course as possible, and with sufficient notice (at least two to three weeks), so that we can work together to make alternate arrangements.

Frequently Asked Questions (FAQ)

How is this course different from ECO375?

Both courses constitute an introduction to econometrics beyond the 2nd year empirical methods that you have already seen, and both courses discuss how different quantitative methods (all based on regressions) can be used to answer research questions. Therefore, the material has some similarities. But the focus of ECO375 is more on understanding the statistical properties of these methods, while ECO372 focusses on their application, and understanding in what contexts these methods can be fruitfully used (or not). As such ECO375 has more algebra and proofs, while reasoning and logic plays an important role in ECO372. ECO375 is recommended if you intend on carrying on postgraduate studies, ECO372 less so.

Is there a lot to memorize?

This course is a methodological course: you will learn about new methods, how they have been used by researchers, and how *you* can use them. The same method applied in different contexts will face different issues and yield different results. Therefore the course emphasis is less on memorization and more on understanding these methods. I test your understanding by asking you to apply the methods to real data (assignments), and by asking reasoning questions about the methods in the quizzes/tests. There are however a few important formulas that you must know and, more importantly, understand.

Is this course math-intensive?

There is *some* algebra, and knowledge of 2nd year statistics/quantitative methods is vital, as basics will not be covered again. However, the focus is not on proofs and calculus *per se*. The course uses equations and formulas (you will see regression equations almost every week of the course), but the focus is on understanding what those formulas can teach us, and how we can interpret them. The course also requires that some familiarity with abstract thinking, and makes heavy use of mathematical notation.

Is this course easy?

The fact that the course is not overly mathematical does not mean it is easy. The course requires good understanding of the methods, in order to apply them in a context that is new to you. In particular this course therefore requires good *reasoning* skills. (The final assessment contains true/false questions in which you must assess the validity of a statement.)

Do I need to buy the textbook?

Yes, we are using the textbook quite heavily (including the mathematical appendices). The textbook has its quirks (and I'll point them out in class), but it is written by two well-respected professors, in a style that is very accessible. It is short, and reads like a novel. It is cheap compared to standard textbooks (around 40CAD), and not heavy. You are free to supplement it with additional readings (eg fom Stock and Watson *Introduction to Econometrics*).

Do I need to buy Stata? Can I use R instead?

Yes you need a working version of Stata (14 or above) for the assignments, and for the tutorials. No, you cannot use R instead. All the statistical analysis in this course is performed using Stata. Stata is a de-facto standard for econometrics, and is widely used by economists. Stata has its quirks, but it has an easier learning curve than R, and the help documentation is consistent and excellent.

Stata comes in different "flavours"; Stata/IC is the cheapest flavour. Although it has limitations compared to the more expensive flavours (hello price discrimination), Stata/IC capacities are well above what is needed for this course.
Stata is also available for free on computers of the Map and Data Library room at Robarts Library.

Can I use Stata 10/11/12/13 instead of the latest version?

All commands we see in the course are the same. However, Stata versions 13 and below are unable to open Stata 14/15/16 datasets. You will not be unable to open most of the files I provide. I advise you against using Stata 13 or below. If you decide to do so anyway, it is your responsibility to find a charitable person that can convert the provided datasets to Stata 13 datasets using `saveold`.

Am I going to become a Stata/programming wizard?

No. This course uses Stata as a *tool* towards data analysis. Doing so requires to learn Stata basics. While I aim to equip you with the basics, and good programming habits, you will not become a Stata or programming expert.

Do I need to already know Stata for this course?

All materials & skills necessary for this course will be provided. Pre-existing knowledge of Stata is not required.

I submitted my assignment only 2 minutes after the deadline, can you waive the penalty?

No. It is your responsibility to allow ample time for submission.

I have another test/assignment due the same week as our assignment/quiz, can I skip/submit late?

See policy about Missed term work.

How do I do well on this course?

Work regularly and use all the help you can.
This course is not easy. It makes heavy use of abstract thinking using mathematical notation, and logical reasoning; and new material arrives fast. To do well, you need to make sure you understand the material well. Working regularly ensures that you digest material as fast as new material arrives. If you let things pile up, you'll only have time for superficial understanding.
I will not be holding your hand; you should take ownership of your own learning. Make sure to use all the help you can, and use all the resources at your disposal: attend all lectures/tutorials, ask questions during office hours or on Piazza, check the textbook, other textbooks if needed, etc.

Is the final exam cumulative?

Yes.

Is there extra work I can do to improve my grade? / Can you raise my course grade by X points? I really need to preserve my GPA.

No, as this would be unfair to other students.