

POL419H1S: Quantitative Methods and Data Analysis

University of Toronto
Winter 2017

Meeting Room:	SS 561
Meeting Time:	Wednesday, 12:00pm – 2:00pm
Instructor:	Kenichi Ariga
Email:	kenichi.ariga@utoronto.ca
Office:	SS 3047
Office Hours:	Monday, 10:00am – 12:00pm

Overview

This is an advanced-level seminar on quantitative empirical research methods for political science for those who have taken POL232, POL242, or equivalent, and have basic understanding of statistical inference and linear regression model.

The primary goals of this class are:

1. To learn the basics of *statistical computing* using R, a freely available, increasingly popular statistical software (<http://www.r-project.org>); and
2. To understand the theoretical foundations, various methods, and applications for *causal inference* in political science research.

Part 1. Statistical Computing

Quantitative political science research requires the use of computers. In the past three decades, the development of affordable yet very powerful personal computers has revolutionized the use of statistical analyses in political science. Various statistical models have been developed and made readily available for researchers. The number of political scientists who have been trained for applying those models has increased, and the volume of applied empirical work using quantitative data and methods has skyrocketed. One of the key driving forces behind these advancements is the wide adaptation of statistical software among political scientists.

In this class you will learn the basics of how to use R to conduct statistical analyses in political science research. Being designed as a political science course rather than a computer programming course, the class will place emphasis on using the program, applying models, and interpreting results rather than on learning how to program. By the end of the semester, you are expected to be able to conduct a basic quantitative empirical analysis using R on your own.

Part 2. Causal Inference in Political Science Research

Those who have taken courses on statistical inference and regression must have heard a mantra that goes “correlation is not causation.” It correctly describes the fact that statistically and substantively significant coefficients in a linear regression model do not necessarily imply a causal relationship of the variables in question. Quantitative empirical evidence in political

science research — many of which has been based on observational data — is to be carefully interpreted with this fact in mind.

Recent development in quantitative empirical analysis on social science research, however, has shifted its attention to the question of when our analysis can provide empirical support to a causal claim. Based on the counterfactual framework, also called the Rubin causal model, specific conditions under which statistical analysis of observational data can provide causal inference have been identified. Moreover, many researchers have also begun to adopt various “design-based” researches — experiments and natural experiments — in which they try to identify a causal relationship mainly from how to design empirical research rather than from how to statistically adjust observed data.

The second part of this class will review these recent attempts of causal inference in political science research. It will be a reading seminar on recent applications in political science research of major research designs and methods for causal inference, such as laboratory experiments, field experiments, survey experiments, matching, natural experiments, instrumental variable analysis, and regression discontinuity design. Students are expected to learn the basic theoretical framework of causal inference and various research designs applied in the current political science research.

Textbook

Thad Dunning. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge: Cambridge University Press.

The textbook is available at the UofT BookStore and will serve as a background reading for the reading seminar in the second part of the class. All other readings on the syllabus will be made available through the class Blackboard site.

Blackboard / Learning Portal

The class Blackboard site (or the Learning Portal: <https://portal.utoronto.ca/>) will be the primary means through which class announcements and assignments will be distributed. Readings and datasets will be made available in the class Blackboard site as well. Its Discussion Board will be the primary medium by which you will ask simple questions about the course materials and get them answered (more on this below). It will be your responsibility to obtain access to the class Blackboard site and regularly check it. There will be an important update to the class Blackboard site at least once a week.

Discussion Board

We will use the Discussion Board in the class Blackboard site as the main medium through which you can ask simple questions regarding class materials and get answers. Given the nature of the course materials, someone else may have the same question as yours and s/he would benefit from your posting the question and getting an answer through the Discussion Board.

You are also encouraged to post an answer to the questions posted by your classmates so that we can maintain a mutually-supporting learning community from which all of you will

benefit. As specified below, your response to your classmate's questions on the Discussion Board will be reflected on your class participation mark.

The instructor will regularly check the Discussion Board and answer questions which have not been adequately addressed by peers. For more complex questions or those that would require an extensive treatment, you are best advised to visit the office hours.

Course Requirements

Your grade will be determined by the following components:

1. R Tutorials & Exercises (Part 1): 15%

R Tutorials:	5%
R Exercises:	10%

In the first part of the class, a series of self-study R tutorials will be assigned every week. You should read and practice what's in them on R. There will also be R exercises based on these tutorials every week. You are required to do all these exercises. R tutorials and exercises assigned in that week are all due by the beginning of the next class.

R tutorials will be graded on a pass/fail basis. If you complete each R tutorial before its due, you will be given a pass — you will receive a full credit for this tutorial. When you complete each R tutorial, you should report it through the class Blackboard site. If you complete an R tutorial after its due, you will be given a marginal pass and earn half the credit for this tutorial. If you don't complete it, you will be given a fail and no credit. In total, R tutorials count toward 5% of your final mark. Completion of each R tutorial will be weighted equally.

R exercises will also be graded on a pass/fail basis. You should show the results of your R exercise to the instructor in the class in which the exercise is due. The instructor will give a quick scan to determine whether you have shown reasonable effort to complete all problems in the exercise. If it is determined that you gave it reasonable efforts to complete all exercise problems, you will be given a full credit for this exercise. If you do not show a reasonable amount of work, however, your exercise will be given a fail or a marginal pass. You will receive no credit in the former case and will receive half the credit in the latter. An example R script for each exercise will be posted on the class Blackboard site after its due. In total, R exercises will count toward 10% of your final mark. Completion of each R exercise will be weighted equally.

2. Two Data Analysis Essays (Part 1): 55%

Data Analysis Essay 1:	30%	due Wednesday, Mar. 1st, 11:59pm
Data Analysis Essay 2:	25%	due Wednesday, Mar. 22nd, 11:59pm

In the first part of the class, there will be two essay assignments. In both assignments, you will conduct data analysis using R and write a short essay on the results.

3. Seminar Presentation (Part 2): 15%

In the second part of the class, we will have reading seminars. Each week, we will read and discuss several articles published in leading political science journals. For each article, one student will play the role of an author, make a presentation on the main findings of the article, and participate in the discussion from the author's perspective. Another student will play a role of a discussant and make a presentation to offer constructive critiques to the article and raise important discussion questions. Together, they will lead the discussion on each article. You will be assigned to either role at least once (the exact number of occasions will be decided and adjusted based on the number of students taking the class). This seminar presentation will count for 15% of your final mark.

4. Critiques and Discussion Questions on the Assigned Readings (Part 2): 5%

For the reading seminar in the second part of the class, you are required to post a short paragraph of critiques and discussion questions to each of the assigned readings of the week on the Discussion Board of the class Blackboard site before noon on Tuesday. The post will be used as a reference for our in-class discussions.

Critiques and discussion questions to each article will be graded on a pass/fail basis. If you post them by their due (Tuesday noon), you will be given a pass (a full credit). If you post them after their due, you will be given a marginal pass (half the credit). If you don't post them, you will be given a fail (no credit). In total, the Discussion Board post of critiques and discussion questions will count toward 5% of your final mark. Each article will be weighted equally.

5. Class Participation (both Part 1 and 2): 10%

Your class participation marks will be determined by the following items:

a. Lectures, Discussion Board Answers, and Collaboration: 9%

This component of the participation mark will be determined by:

- 1) whether you are actively engaged in class lab sessions and seminar discussions (i.e., regularly attend the class, actively raise questions, participate in class discussion, help your classmates);
- 2) how often and well you respond to your classmates' questions on the class Discussion Board; and
- 3) whether you collaborate with your classmates in essay assignments (more on this below).

d. Feedback Survey: 1%

There will be an anonymous online feedback survey on the class through the Blackboard at the end of the semester. Your participation in the survey will count toward 1% of your final mark.

Group Work and Collaboration

Group work and collaboration is encouraged in this class. Given somewhat technical nature of the class materials, it is essential to have an opportunity to discuss with your classmates the concepts and methods you learn in class and how to apply them. Everyone has different strengths and weaknesses in their understanding of materials and learning style. Through working together, you are expected to facilitate learning for each other and deepen your understanding of the materials, which would be difficult if you worked alone.

Collaboration in a team is encouraged in this class so much so that the submission by a team is allowed for the data analysis essay assignments. At most three students may participate in one group. If you submit your assignment as a group, everyone in the group will receive the same mark for that assignment. The group submission of the essay assignments will also be taken into consideration when the “Lectures, Discussion Board Answers, and Collaboration” component of your class participation mark is determined.

Collaboration in a team of multiple scholars is a norm for contemporary social science research in general, and quantitative empirical political science research in particular. As a course on the methods of such research, this class will provide you with an opportunity to practice scholarly collaboration by allowing the group submission of the assignments.

Turnitin

Normally, students will be required to submit their course essays to Turnitin.com for a 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.

Students who wish to not use Turnitin.com may make an alternative arrangement with the instructor. They will need to let the instructor know well before the deadline of the assignment and ask for an alternative way to submit the essay. They will be required to save every version/draft of their essay electronically, and submit all of them at the time they submit the essay. They will also be required to hand in all notes, outlines, and bibliographic research at the same time.

Late Penalties and Extension

All work is late if submitted after the date and time specified as due. To ensure fairness, the late-penalty policy specified below will be strictly enforced. Conflict with other class’s assignment/exam schedule, leaving for a non-academic trip, or vacation is not an acceptable reason to miss the assignments or request an extension.

Extension for the data analysis essay assignments may be granted only when there is a legitimate reason, such as an unforeseeable medical emergency and an accessibility issue, and there is an acceptable official documentation, which verifies the specific reason given, such as the UofT Verification of Student Illness or Injury form, the Accessibility Services Letter, and the College Registrar’s Letter. Students who know in advance they will need an extension for

a legitimate reason should contact the instructor as early as possible before the deadline. Those who missed the deadline for a legitimate, unforeseeable reason should contact the instructor as soon as possible and no later than one week after returning to class.

Data analysis essay assignments handed in late will result in a penalty of 2-percentage-point reduction per day (e.g., from 72% to 70%). Submitting an essay within 24 hours from the due date and time will be considered one day late; submitting after 24 hours but before 48 hours will be two days late, and so forth.

Since the Turnitin is used to submit essay assignments, your submission must be accepted by Turnitin before the due date and time. Note that the date and time recorded in Turnitin will be your submission date and time. If this is after the due date and time, then your submission will be considered late. In other words, completing your essay and start uploading it to Turnitin before the due date and time is not enough. Your submission to Turnitin must be complete before the due date and time.

Also note that your submission is incomplete unless you receive a Turnitin submission ID. It is your responsibility to make sure that you received a submission ID before the due date and time.

Computer-related problems, such as the crash of your computer, a slow Internet connection, and an occasional slow response of the server, will not be considered as an acceptable reason to request for extension or waiver of a late penalty. Also sending your assignment to the instructor via email will not be considered as a submission. For these reasons, I strongly suggest you avoid a last-minute completion or submission of assignments. I also suggest you frequently take a backup of the electronic files of your draft essay in an electronic storage other than your computer.¹

Grade Appeals

Grade appeals must be received within two weeks from when the grade is assigned. When you appeal your grade, you are required to submit a brief documentation substantiating why you believe the grade is not appropriate.

Office Hours

You are welcome to visit during the instructor's office hours, which will be held during the time and date specified at the beginning of the syllabus, if you have any questions on the class materials.

Email Policy

If you have questions of personal nature (e.g., accessibility, deadline extension for legitimate reasons), you may email the instructor and expect a response within two working days. Please start the subject heading of your email with "POL419:..." I will not respond to, however, any

¹ For example, you may use cloud storage spaces, such as DropBox, Google Drive, and iClouds. Or you may send your draft to your UofT email address so that your draft file will be stored in your mailbox of the UofT server.

questions over email that are of substantive nature concerning the class materials. You will need to post those questions on the Discussion Board or visit office hours.

Please note that I will not be able to answer emails or Discussion Board questions during weekends.

In the case of your questions of substantive nature on the Discussion Board or those of personal nature over email not answered within two working days (excluding weekends), send me an email to let me know they have not been addressed. Please include “POL419: Unanswered Question” in the subject heading of your email.

Accessibility

The University of Toronto is committed to accessibility. If you require accommodation for a disability, or have any accessibility concerns about the course, the classroom or course materials, please contact Accessibility Services at www.accessibility.utoronto.ca, accessibility.services@utoronto.ca, or (416) 978-8060 as soon as possible.

Academic Integrity

Academic integrity is fundamental to learning and scholarship at the University of Toronto. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the U of T degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves.

You are expected to be familiar with the Code of Behaviour on Academic Matters, available at <http://www.artsci.utoronto.ca/osai/students>, which is the rule book for academic behaviour at the U of T. Potential offenses include, but are not limited to, plagiarism, cheating on tests and exams, fraudulent medical documentation and improper collaboration on marked work.

For specific examples of the potential academic offences, please read *What is Academic Misconduct* (<http://www.artsci.utoronto.ca/osai/The-rules/what-is-academic-misconduct>) at the Office of Student Academic Integrity’s website. Please note that, as stated in this site, “(n)ot knowing the University’s expectations is not an excuse.” Under the Code, “the offense shall likewise be deemed to have been committed if the person ought reasonably to have known.” (*Code of Behaviour on Academic Matters*, web version, p.2)

For further clarification and information on plagiarism, please see *Writing at the University of Toronto* (<http://www.writing.utoronto.ca/advice/using-sources/>).

The University of Toronto treats cases of academic misconduct very seriously. All suspected cases of academic dishonesty will be investigated following the procedures outlined in the Code. The consequences for academic misconduct can be severe, including a failure in the course and a notation on your transcript. If you have any questions about what is or is not permitted in this course, do not hesitate to contact the instructor.

Class Schedule

Part I: Introduction to Statistical Computing (Lab Sessions)

The self-study R tutorials written by the instructor will be used in this part of the class.

Jan. 11. Introduction

- Introduction to R & RStudio.
- Very basic operations of R.
- Descriptive Statistics & Visualization in R.
- R Tutorials 0 through 4.

Jan. 18. Data Analysis using R (1)

- Linear Regression Analysis and Data Management in R.
- R Tutorials 5 through 6.

Jan. 25. Data Analysis using R (2)

- Linear Regression Analysis and Data Management in R, continued.
- R Tutorials 7 through 9.

Feb. 1. Data Analysis using R (2)

- Linear Regression Analysis and Data Management in R, continued.
- R Tutorials 10 and 11.

Feb. 8. Data Analysis Essay 1 Consultation

- Q & A on your project for Data Analysis Essay 1.

Feb. 15. Statistical Simulation

- Monte Carlo Simulation.
- R Tutorial 12.

Feb. 22. No Class (Reading Week)

Mar. 1. Part I Wrap-Up

Data Analysis Essay 1 Due: Mar. 1 (Wed.), 11:59pm

Data Analysis Essay 2 Due: Mar. 22 (Wed.), 11:59pm

Part II: Causal Inference in Political Science Research (Reading Seminar)

Mar. 8, 15, 22, 29 and Apr. 5.

The reading list below is tentative. It will be finalized by the first lecture after the reading week.

Laboratory Experiment

Fridkin, Kim L., Patrick J. Kenney, Sarah Allen Gershon, Karen Shafer, and Gina Serignese Weedall C. 2007. "Capturing the Power of a Campaign Event: The 2004 Presidential Debate in Tempe." *Journal of Politics* 69(3).

Field Experiment

Malesky, Edmund, Paul Schuler, and Anh Tran. 2012. "The Adverse Effects of Sunshine: A Field Experiment on Legislative Transparency in an Authoritarian Assembly." *American Political Science Review* 106(4).

Standard Survey Experiment

Hainmueller, Jens, and Michael J. Hiscox. 2010. "Attitudes toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment." *American Political Science Review* 104(1).

Endorsement Experiment

Lyall, Jason, Graeme Blair, and Kosuke Imai. 2013. "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107(4).

Matching

Cindy D. Kam and Carl L. Palmer. 2008. "Reconsidering the Effects of Education on Political Participation." *Journal of Politics* 70(3).

Henderson, John and Sara Chatfield. 2011. "Who Matches? Propensity Scores and Bias in the Causal Effects of Education on Participation." *Journal of Politics* 73(3).

Mayer, Alexander K. 2011. "Does Education Increase Political Participation?" *Journal of Politics* 73(3).

Cindy D. Kam and Carl L. Palmer. 2011. "Rejoinder: Reinvestigating the Causal Relationship between Higher Education and Political Participation." *Journal of Politics* 73(3).

Weighting

Truex, Rory. 2014. "The Returns to Office in a 'Rubber Stamp' Parliament." *American Political Science Review* 108(2).

Natural Experiment

Bhavnani, Rikhil R. 2009. "Do Electoral Quotas Work After They Are Withdrawn? Evidence from a Natural Experiment in India." *American Political Science Review* 103(1).

Instrumental Variable

Hansford, Thomas G., and Brad T. Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104(2).

Regression Discontinuity

Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review* 109(1).

Syllabus Change Policy

The policies and contents of this syllabus may be changed by the instructor with advanced notice. If any, such a change will be announced during lectures.