POL419H1S/POL2519H1S: Quantitative Methods and Data Analysis
Part I: Basics of Statistical Computing and Statistical Analysis
Part II: Causal Inference for Observational Data

University of Toronto
Winter 2015

Meeting Room:	SS 561
Meeting Time:	Friday, 10:00am – 12:00pm

Instructor:	Kenichi Ariga
Email:	kenichi.ariga@utoronto.ca
Office:	SS 3047
Office Hours:	Wednesday, 10:00am – 11:50am

Additional Office Hours by Dr. Hanil Chang

Email:	hanil.chang@utoronto.ca
Office:	SS 3045
Office Hours:	Tuesday, 14:00-16:00 Thursday, 15:00-17:00

Overview
This is an advanced-level class on quantitative empirical research methods for political science for those who have taken POL242, POL322, POL2504 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; and
2. To understand the theoretical foundations and various methods for causal inference for observational studies.

1. Statistical Computing
Quantitative social science research requires the use of computers. In the past two decades, the development of affordable yet very powerful personal computers has revolutionized the use of statistical analyses in social sciences. Various statistical models have been developed and made readily available for researchers. The number of social 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 social scientists.

In this class you will learn the basics of how to use R to conduct statistical analyses in political science research. This class being designated as a political science course rather than a computer programming course, the emphasis will be placed on using the program, applying models, and interpreting results rather than on learning how to program.

2. Causal Inference for Observational Studies

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 social 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 — often called the Neyman or 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 in which they try to identify a causal relationship mainly from how to design empirical research (e.g., natural experiments) in addition to identifying a causal relationship from statistical adjustment to the observed data.

The second part of this class will study the theoretical foundations and various statistical methods of these recent attempts of causal inference in social science research based on observational data.

Textbook

Joshua D. Angrist and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton University Press.

This textbook will be used for the second part of the class on causal inference for observational studies. Other required readings will be made available through the class Blackboard site.

Computer Software

Quantitative social science research requires the use of computers. In this class, you will use a software package called R, which is free to download at http://www.r-project.org and is getting popular among many social scientists. By the end of the semester, you are expected to be able to conduct a basic quantitative empirical analysis using R on your own.

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 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.

Course Requirements

Your grade will be determined by the following components: 1) homework assignments and 2) class participation.

1) Homework assignments: 80% (Each homework assignment will be weighted equally.)
2) Class Participation and Collaboration: 20% (See below for more detail.)

1) Homework Assignments: There will be at most six homework assignments (actual number will vary depending on the progress of the class), which would take different formats depending on the class materials covered each week. Many of them will involve data analysis and interpretation using R. For these assignments, you will be encouraged to work and submit the assignments in a small group up to three individuals. If you submit your work as a group, everyone in the group will receive the same mark and will also receive a small collaboration credit as part of the class participation mark (see below for more detail). These homework assignments will count 80% toward your final mark, and each of the assignments will be weighted equally.

2) Class Participation and Collaboration: Your class participation grade will be determined by the following three components:

a) Whether you are actively engaged in class lectures (i.e., regularly attend the class, actively raise questions, and participate in class discussion) (10%),
b) How often and well you respond to your classmates' questions on the class Discussion Board (5%, both quantity and quality matter), and
c) Whether you collaborate with your classmates in homework assignments (5%, each assignment weighted equally).

Group Work and Collaboration

Group work and collaboration is encouraged for this class. Given somewhat technical and advanced 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. In fact, 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.

For the assignments of this class, collaboration in a team is encouraged so much so that the submission by a team is allowed. Moreover, if you submit the assignment in a group, you will also earn a small credit toward your class participation and collaboration mark, taking into consideration the coordination cost among group members. Please note that at most three students may participate in one group.

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 the 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
All work is late if submitted after the date and time specified as the due date. Any assignments handed in late will result in a penalty of 20 percentage points reduction per day (e.g., from 90% to 70%). Submitting assignments 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. Assignments handed in more than five calendar days late will receive a zero grade.

Extension
Extension 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 Medical Certificate, the College Registrar's Letter, and the Accessibility Services 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.

Grade Appeals
Grade appeals must be received within one month from when the grade is assigned. When you appeal your grade, you are required to submit a 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. I will maintain a sign-up sheet for my office hours online. Please sign up for a 10-minute block on this sign-up sheet. More details about the sign-up sheet will be posted on the class Blackboard site.

There will also be additional office hours, held by Dr. Hanil Chang, a postdoctoral fellow at the Department of Political Science, during the time and date specified at the beginning of the syllabus. He is also available for appointments upon email request (hanil.chang@utoronto.ca).

Discussion Board
We will use the Discussion Board in the class Blackboard site as the main medium through which you can ask 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 stated above, your response to your classmate's questions on the Discussion Board will be reflected on your class participation and collaboration mark.

The instructor will regularly check the Discussion Board (once on Mondays, Wednesdays and Fridays) 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.

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 "POL419/2519:..." I will not respond to, however, any questions over email that are of substantive nature concerning the class materials. You will need to post those questions on the Discussion Board.

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 "POL419/2519: 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 (416) 978-8060 or www.accessibility.utoronto.ca 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 Topics

The following is the list of topics to be covered throughout the semester. The dates for each topic are not given to allow some flexibility to adjust our course schedule to our needs and pace. More details on each class session and the reading assignments will be announced as we proceed.

During the semester, some assigned readings may be replaced by others. If these are to happen, you will be given an advance notice during lectures.

Part I: Basics of Statistical Computing and Statistical Analysis

1. R Boot Camp: Very Basics of Statistical Computing

- Basic use of R
- Descriptive statistics, such as histograms, boxplots, and scatterplots
- Linear regression analysis

2. Statistical Inference Review

- Random variables and probability distribution
- Expectations of random variables
- Law of large numbers
- Central limit theorem
- Point estimation, confidence intervals, and hypothesis testing
- ➢ Review those topics in the textbook you used in the prerequisite classes.

3. Linear Regression Review

- Conditional probability distribution and conditional expectations
- OLS estimator
- Classical normal linear regression model
- Classical linear regression model
- Modern formulation of linear regression model
- ➢ Review those topics in the textbook you used in the prerequisite classes.

4. Logit/Probit Model for the Binary Dependent Variable

- Maximum likelihood estimation

- Logit/Probit model
- Review those topics in the textbook you used in the prerequisite classes.

Part II: Causal Inference for Observational Studies

5. Theoretical Foundation for Causal Inference: Counterfactual Framework

Required Readings:

- Angrist and Pischke, Chapters 1-2.
- Thad Dunning. 2014. *Natural Experiments in the Social Sciences: A Deign-Based Approach*. Cambridge University Press. Chapter 5.1

Recommended Readings:

- Alan S. Gerber and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. Norton. Chapter 2
- Stephen L. Morgan and Christopher Winship. 2007. *Counterfactuals and Causal Inference*. Cambridge University Press. Chapter 2.

6. Matching 1: Propensity Score Matching

Required Readings:

- Angrist and Pischke, Chapter 3.2-3.3
- Cindy D. Kam and Carl L. Palmer. 2008. "Reconsidering the Effects of Education on Political Participation." *Journal of Politics* 70 (3): 612-631.
- Jasjeet S. Sekhon. 2011. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* 42 (7).

7. Matching 2: Genetic Matching

Required Readings:

- Reread Sekhon 2011 in the previous topic.
- Alexander K. Mayer. 2011. "Does Education Increase Political Participation?" *Journal of Politics* 73 (3): 633-645.
- John Henderson and Sara Chatfield. 2011. "Who Matches? Propensity Scores and Bias in the Causal Effects of Education on Participation." *Journal of Politics* 73 (3): 646-658.
- Cindy D. Kam and Carl L. Palmer. 2011. "Rejoinder: Reinvestigating the Causal Relationship between Higher Education and Political Participation." *Journal of Politics* 70 (3): 612-631.

Recommended Readings:

- Alexis Diamond and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95 (3): 932-945.
- Walter R. Mebane, Jr. and Jasjeet S. Sekhon. 2011. "Genetic Optimization Using Derivatives: The rgenoud Package for R." *Journal of Statistical Software* 42 (11).

8. Regression Discontinuity

- Angrist and Pischke, Chapter 6.1
- Thad Dunning. 2014. *Natural Experiments in the Social Sciences: A Deign-Based Approach*. Cambridge University Press. Chapters 3 and 5.2

- Taylor C. Boas, F. Daniel Hidalgo, and Neal P. Richardson. 2014. "The Spoils of Victory: Campaign Donations and Government Contracts in Brazil." *Journal of Politics* 76(2).

9. Instrumental Variable

- Angrist and Pischke, Chapters 4.1, 4.4, 4.5
- Thad Dunning. 2014. *Natural Experiments in the Social Sciences: A Deign-Based Approach*. Cambridge University Press. Chapters 4 and 5.3
- Kristopher W. Ramsay. 2011. "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy." *International Organization* 65.
- Stephen Ansolabehere, and James M. Snyder, Jr. 2004. "Using Term Limits to Estimate Incumbency Advantages When Officeholders Retire Strategically." *Legislative Studies Quarterly* 29 (4): 487-515.

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.