POLS 642: Intermediate Analysis of Political Data

Northern Illinois University

Spring 2024

Tuesdays and Thursdays 2-3:15

Instructor: Dr. Colin Kuehl
Email: ckuehl@niu.edu
Office Hours: Tuesday & Thursday 1230-130
Office: Zulauf 410 and Zoom

Course Description: This is the second course in the NIU Political Science Department's quantitative methods sequence. It is taken by most graduate students in the department who wish to satisfy the methodological requirements through a balance of quantitative and qualitative methods courses. It is taken by all students seeking to eventually take more advanced courses in quantitative methods outside the department. The course focuses on linear regression models with an emphasis on the concepts of probability and inference, the assumptions of linear regression estimation, and the application of linear models to research questions. Additional quantitative technique will also be introduced. Students shouldleave the course with the ability to engage with quantitative empirical research, to employ basic OLS and logit models in their own research, and to further their statistical skills through either more advanced coursework or self-learning. In order to accomplish these goals, Students will continue to use the R programming language(more below) and will conduct original quantitative analysis using existing data.

Preparation: This course builds heavily off of the skills and concepts learned in POLS 641. Some of you have taken the course recently and for others it has been a couple years. We will spend some time reviewing basic concepts, but will quickly build on the foundation I expect you to know from the previous course. As with 641, you will not need a strong mathematical background to succeed in the course. We fill focus on the practical applications of the models and as necessay open up the hood

Course Meetings: We will meet twice per week. The first, on Tuesday, will be a combination of lecture and discussion on fundamental concepts, such as research design, probability theory, etc. Please make sure all course readings are completed before our Tuesday meeting. The second meeting, on Thursdays, will focus on application and coding in R. Meetings will be primarily be held in person, however depending on scheduling issues and student interest we may move some sessions online.

Readings We have one assigned text for the course, but will rely heavily on two others:

- Michael A. Bailey, *real stats: Using Econometrics for Political Science and Public Policy*, Oxford. (second edition, first edition is OK)

  This is the same book used in 641

Available for free online, this book offers helpful R code to implement all of the analyses, tables, and figures, presented in the Bailey textbook:

- Tony Corelli *An R Companion to Real Econometrics.*

We also will rely heavily this book for more detailed explanations. It is also available for free online.

- Andrew Gelman, Jennifer Hill, and Aki Vehtari *Regression and Other Stories*

In addition to the required text, there are a number of previously mentioned recommended books.

"Pop" data science books:

- Carl Bergstrom & Jevin West *Calling Bullshit*
- Nate Silver *The Signal and the Noise*
- Charles Wheelan *Naked Statistics*
- Tim Harford *The Data Detective*

Basic introductions

- Neil Salkind *Statistics for People who (think they) Hate Statistics*
- Scott Cunningham *Causal Inference Mixtape*
- Gary Klass *Just Plain Data Analysis*

More advanced texts for students looking to go beyond what is taught in the course:

- Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchal Models*
- Damodar N. Gujarati *Essentials of Econometrics*
- Jeffrey Wooldridge*Introductory Econometrics*
- Elena Llaudet and Kosuke Imai *Data Analysis for Social Science*
- Jenine Harris *Statistics with R*
- Guido Imbens and Donald Rubin *Causal Inference for Statistics, Social, and Biomedical Sciences*

Additional assigned readings will be found on the course blackboard page. The blackboard page also contains links to helpful websites and online books, especially for R help.

Programming in R As part of this course, we will continue to use the statistical programming language R. It is assumed that students have developed a solid introduction to R in POLS 641 and we will build on that foundation. I will continue to teach R skills in class, we will do some of DataCamp exercises, and I have given you the Carilli R handbook that accompanies the Bailey textbook. But I will ask you to increasingly figure things out on your own. Part of the skillset being developed in this class is the ability to teach yourself going forward, both with regards to statistics as well as to programming.

Evaluation: Final Project (40%), Problem Sets (20%), Replication Project (20%), Teaching Demo (5%) Datacamp (5%), Participation (10%),

- **Final Project:** Your final project will be an independent quantitative analysis of a political science topic of your choosing. You will obtain a dataset, formulate a hypothesis, conduct basic exploratory data analysis and perform, as well as correctly interpret, an advanced model appropriate for your dataset. You will present your results to the class at the end of the semester and write up what you found in the form of the empirical section of journal article. An assignment sheet will be provided and we will discuss details in the following weeks.

  You will present your findings to the class at the end of the semester. Your final project will be due **May 8th**.

- **Problem Sets:** Throughout the semester you will be expected to complete a number of problem sets. These exercises will give you a chance to apply the ideas and concepts using the R coding we are learning. Problem sets will be distributed regularly starting in week 3 and will be due on Wednesday at midnight of the second week.

  I strongly encourage you to form small study groups to work through the problem sets. However, the write-ups *must be your own*. This means that you can talk through how to solve a problem with someone else, have them check your code for errors, etc. but you must then on your own (later in the day, in silence) put the solution down in writing. No late assignments will be accepted without prior permission. Extensions should be arranged at least twenty-four hours in advance.

- **Replication or Experimental Design Project:** Students will have the option of replicating an existing scholarly result or can choose to design a novel survey experiment. The choice depends on which method will be most useful for their own scholarly careers. For replications you will be expected to find previously published article, obtain the data, and then write the R code to replicate as closely as possible the published figures and model results. For experiments you will be expected to essentially create the materials to preregister a survey experiment. This includes writing a treatment, writing questions for outcome variables and demographics, and a basic analysis plan. More details will be discussed in class.

  Students have the option to complete this project in pairs, however expectations will be increased proportionally. Choose wisely.

- **R Teaching Demonstration** Each student will record a video tutorial demonstrating a fundamental R skill. This will be an opportunity reacquant yourself with R and demonstrate your abilities as a teacher. The video plus practice code and data will be submitted. Videos will posted on youtube as a resource for future students. A sign-up sheet will distributed in week one.

- **Datacamp:** As a member of the class you will a complementary six-month membership. These courses are intended to give you some additional practice and exposure to differing techniques as a supplement to what is taught in the course. For this course the choice of datacamp course will be up to you. Twice during the semester you will submit screenshots of two courses that you have completed.

- **Participation:** Class time will be divided between discussion and lecture of the various statistical concepts and techniques(on Tuesdays) and time spent learning and practicing in the R programming language(on Thursdays). My expectation is that this will be a true discussion and students will ask questions when they are confused or having difficulties. Your active participation is especially important given the wide range of backgrounds and small class size. Speaking the language of methodology is an essential component of your graduate training.

Due Dates Unless specified otherwise all course assignments will be due Wednesday at midnight

Date	Whats Due?
1/17	
1/24	Datacamps 1
1/31	Teaching Demo
2/7	Pset #1
2/14	Replication Project Proposal
2/21	Pset #2
2/28	Final Project: Abstract and Hypothesis
3/6	Pset #3
3/13	*Spring Break*
3/20	Datacamps 2
3/27	Replication Project
4/3	
4/10	Final Project Data and Initial Analysis
4/17	Pset #4
4/24	
5/1	Research Presentations
5/8	Final Project Due

Course Policies:

- Attendance: Attendance is mandatory without prior arrangement. If circumstances prevent you from attending please let me know by email prior to the beginning of class.. You will be responsible for covering the material you missed.
- Academic Honesty: Cheating will not be tolerated. All students will be held to the highest standards of NIU's student code of conduct. All cases will be referred to campus authorities. As noted above, helping each other will be key to your sucess in this class, however the work you turn in must be your own.
- A note on AI: The use of intelligence, such as ChatGPT and other language models is allowed in this course. I believe it may become a valuable tool in your academic workflow, particularly when coding. However, any writing done for the course is expected to be yours own. Solely relying on AI-generated content without proper attribution is still plagiarism. If using these tools you must cite them appropriately. There are a myriad of ways to productively use this technology, but please do so responsibly and with a critical eye.
- Disability Services: If you need an accommodation for this class, please contact the Disability Resource Center as soon as possible. The DRC coordinates accommodations for students with

disabilities. It is located on the 4th floor of the Health Services Building, and can be reached at 815-753-1303 or drc@niu.edu.Also, please contact me privately as soon as possible so we can discuss your accommodations.

- Email: I am available through email M-F 8 am to 5 pm. I will respond to all emails within 24 hours during these times. If I forget, please send me a reminder.
- Late Assignments: Late assignments will not be accepted without prior approval of the instructor. If you need more time, just ask.
- Office Hours: I recommend you come to office hours early and often. This will give me a chance to know you, your projects, etc. and give us a chance to talk about things beyond this course
- Mental Health: Graduate school is challenging. We all will handle this uncertainty and stress in different ways. Please take the time for self-care and if you need extensions for mental health reasons do not hesitate to ask. If you or somebody you know is struggling with anxiety or other issues, do not hesitate to reach out. Resources available include the DRC, Student Counseling Services or call 815-306-2777.
- Health Precautions: Masks are not required for this class (subject to changes in protocols from the univeristy). Please be respectful of others decisions. Above all, please be considerate of your fellow students. If you are feeling sick stay home. If you might be sick wear a mask. At all times, please do your part to keep everyone healthy.
- Gender Identity: As a faculty member, I am committed to using your proper name and pronouns. We will take time during our first class together to do introductions, at which point you can share with all members of the course what name and pronouns you use, as you are comfortable. Additionally, if these change at any point during the semester, please let me know and we can develop a plan to share this information with others in a way that is safe for you.
- Land Acknowledgement Statement: Northern Illinois University operates and is built on the traditional lands of the Oceti Sakowin (Sioux), Miami, Bodewadmiakiwen (Potawatomi), Sauk and Meskwaki, and Peoria. These lands are subject to Cession 50 and 148, though their terms have been consistently violated. We seek to acknowledge this land and these peoples in order to honor the legacies, struggles and current existence of Indigenous peoples; situate ourselves within settler-colonial projects; disrupt the erasure of Indigenous peoples; and begin/continue the work of collectively learning and fulfilling our obligations, for those of us who are uninvited guests on Indigenous peoples' lands

Tentative Course Outline:

The tentative schedule for the course is below. Some topics will take just a week, others more. Given our small size and wide variety of backgrounds, we will be flexible in how fast we move. If we move quickly, there are a number of optional topics we can cover at the end of the semester. Treat this as a preliminary overview, and not a complete roadmap. Reading assignments (beyond your book chapters) for the following week will be provided at the end of each class and posted on Blackboard.

Week	Date	Course Topic
1	16-Jan	Introductions
2	23-Jan	Fundamentals Review(Probability, Measurement, etc)
3	30-Jan	Causal Inference Using Potential Outcomes Framework
4	6-Feb	Regression Review
5	13-Feb	Variables: Dummys, Categorical, Interactions
6	20-Feb	Model Specification
7	27-Feb	Binary DVs, Logistic Regression
8	5-Mar	Time-Series and Panel Data
9	12-Mar	Spring Break
10	19-Mar	More Time: Binary Time-Series Cross-Sectional
11	26-Mar	Diff in Diff and Regression Discontinuity
12	2-Apr	Scale Construction, Validity, Reliability *(MPSA)*
13	9-Apr	Multilevel Models
14	16-Apr	Network Analysis Basics
15	23-Apr	Class Choice
16	30-Apr	Research Presentations
17	7-May	Projects Due

Course Readings and Due Dates:

Week 01, 01/16 : Introductions
Readings:

- Chaudhuri, A., Iversen, V., Jensenius, F. R., & Maitra, P. (2024). "Time in Office and the Changing Gender Gap in Dishonesty: Evidence from Local Politics in India. *American Journal of Political Science*, 68(1), 106–122. https://doi.org/10.1111/ajps.12733
- RS: Chs. 1-2

Recommended:

-

Week 02, 01/23 : Reviewing the Fundamentals of Quantitative Analysis
Readings:

- McCauley, Adam, & Reggeri, Andrea. (2020). "From Questions and Puzzles to Research Project." In Curini, L., & Franzese, R. *The SAGE Handbook of Research Methods in Political Science and International Relations*.
- Other Stories: Chs. 1-2

Recommended:

- Ch 3 Probability *OpenIntro Statistics*

Week 03, 01/30 : Causal Inference Using the Potential Outcomes Framework
Readings:

- Angrist & Pischke *Mostly Harmeless Econometrics* Chp 1-2
- Pearce, Neil and Debbie A. Lawlor. 2016. "Causal inference–so much more than statistics." *International Journal of Epidemiology*:
- Broockman, David and Joshua Kalla "Durably reducing transphobia: A field experiment on door-to-door canvassing" *Science* 352,220-224(2016).

Recommended:

-
- Scott Cunningham *Causal Inference Mixtape*

Due: Datacamp - R Basics

Week 04, 02/06 : Regression Review
Readings:

- RS: Chs. 3, 14.1, 14.2
- Other Stories: Chs. 6 & 7

Recommended:

- Miller *How to Read a Regression Table*

Week 05, 02/13 : Variables: Dummys, Categorical, Interactions
Readings:

- RS: Ch. 6
- Other Stories: Ch 10
- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." Political Analysis 14: 63–82.

Recommended:

- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." Political Analysis 27:163-192.

Week 06, 02/20 : Model Specification
Readings:

- RS: Ch. 7
- Other Stories: Ch 11
- Meuleman, Bart, Geert Loosveldt & Viktor Emonds "Regression Analysis: Assumptions and Diagnostics" in *The SAGE Handbook of Regression Analysis and Causal Inference*
- Healy, Data Visualization, Ch. 6

Recommended:

- Carilli: Ch. 9
- King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." American Journal of Political Science 44(2): 347–61.
- Other Stories: Chs. 6, 12, Appendix B

Week 07, 02/27 : Binary Dependent Variables
Readings:

- RS: Ch. 12
- Other Stories: Chs. 13, 14

Recommended:

- Carilli: Ch. 14

Week 08, 03/05 : Time Series and Panel Data
Readings:

- RS: Chs. 13, 15
- Troeger, V. (2020). "Time-Series-Cross-Section Analysis" In Curini, L., & Franzese, R. *The SAGE Handbook of Research Methods in Political Science and International Relations.*

Recommended:

- Linn, S., & Webb, C. (2020). "A principled approach to time series analysis." In Curini, L., & Franzese, R. *The SAGE Handbook of Research Methods in Political Science and International Relations.*

Week 09, 03/12 : *Spring Break*

Week 10, 03/19 : More Time Series and Differences-in-Differences
Readings:

- RS: Ch. 8
- Keele, L. (2020). "Differences-in-differences: neither natural nor an experiment" In Curini, L., & Franzese, R. *The SAGE Handbook of Research Methods in Political Science and International Relations.*

Recommended:

- Beck, Nathaniel, Katz, Jonathan N., & Richard Tucker (1998). "Taking Time Seriously in Binary Time-Series–Cross-Section Analysis." *American Journal of Political Science* 42(4):1260-1288

Week 11, 03/26 : Instrumental Variables and Regression Discontinuity
Readings:

- RS: Ch. 9
- RS: Ch. 11

Recommended:

- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Cambridge University Press.

Due: Problem Set #4

Week 12, 04/02 : Scale Construction
Readings:

- Munck, G., Møller, J., & Skaaning, S. 2020. "Conceptualization and measurement: basic distinctions and guidelines" In Curini, L., & Franzese, R. *The SAGE Handbook of Research Methods in Political Science and International Relations.*
- Carmines, Edward G., and Richard A. Zaller. 1979. Reliability and Validity Assessment. Newbury Park, CA: Sage Publications

Recommended:

-

Note: No class on Thursday for MPSA

Week 13, 04/09 : Network Analysis Basics (Guest Lecture: Dr. Scott Robinson)
Readings:

- TBD

Recommended:

-

Week 14, 04/16 : Multilevel Models
Readings:

- TBD
- Other Stories: Ch 22

Recommended:

-

Week 15, 04/23 : Class Choice: Text as Data & Using AI to help code
Readings:

- Grimmer, Justin. 2015. "We are all social scientists now: How big data, machine learning, and causal inference work together" PS:

Recommended:

- Grimmer, J., Roberts, M.E. and Stewart, B.M., 2022. *Text as data: A new framework for machine learning and the social sciences.* Princeton University Press.
- Roberts, M.E., Stewart, B.M. and Tingley, D., 2019. *Stm: An R package for structural topic models. Journal of statistical software*, 91, pp.1-40.

Week 16, 04/30 : Research Presentations