GOV 391L: Statistical Analysis in Political Science II

Spring 2023

MW 10:00-11:30am, SZB 4.406

Professor: Alison Craig
Email: acraig@austin.utexas.edu
Office Hours: MW 1:00-2:30pm
Office: BAT 4.148

TA: Matt Martin
Email:
Office Hours:
Office:

Course Description

This course will provide students with an understanding of the estimation, diagnostics, and interpretation of statistical models commonly used in political science research. The course begins with an examination of the mechanics of Ordinary Least Squares (OLS) and continues with Generalized Linear Models (GLMs), Maximum Likelihood Estimation (MLE), and an introduction to causal inference for observational data. The course will require a significant time commitment and will move at a rapid pace. In addition to the class meetings, students are expected to keep up with the readings, complete biweekly problem sets, and attend weekly discussion sections that will focus on reviewing material and working through applications in the R statistical computing environment. By the end of the semester, students should be comfortable working with the standard political methodology toolkit.

Course Requirements and Grading

Your grade in this course will be based on four components as follows:

Problem Sets	50%
Midterm Exam	25%
Final Project	25%

There will be five problem sets assigned over the course of the semester. Students are encouraged to work in groups on problem sets and in studying for exams, but all work must be written up individually. All problem sets should be typed up in LaTeX or R Markdown and include a clean copy of any R code used in a fixed width/monospaced font. Problem sets must be submitted via Canvas prior to the start of class on the day the assignment is due. **Leaving the problem sets for the last minute will cause you unnecessary stress and sadness.** Late assignments will only be accepted with prior approval (i.e. reach out to me *before* the deadline) and will receive a ten percentage point deduction for non-emergency situations.

The midterm exam will be an in-class exam on the material covered during the first three units of the course (OLS, Model Specification, and Diagnostics) and is tentatively scheduled for March 8.

In lieu of a final exam, students will complete a final project consisting of a presentation and a paper utilizing the tools you will learn this semester. The final paper will be a complete research paper that includes a well-defined research question, a theory grounded in the research in your field, testable hypotheses, data analysis using a least-squares regression method, and robustness checks. You will work with one of your classmates to either conduct a replication and extension of a previously published paper in political science or write an original research paper using existing data sources. The paper is due April 30.

You will also present your research in the final week of the class in the style of a conference presentation, in which you have 10-12 minutes to present your work and presentations are followed by class discussion. Students will be graded on both their own presentation and their participation in discussions. All students are expected to attend all presentation days to support their classmates and contribute to discussion.

Required Texts

- Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.
- Michael D. Ward and John S. Ahlquist. *Maximum Likelihood for Social Science: Strategies for Analysis*. Cambridge University Press, 2018
- Jeffrey Wooldridge. *Introductory Econometrics: A Modern Approach*. Cengage, 6th edition, 2016

In addition to the required text, you will need R, R Studio, and LaTeX installed on your computer. Journal articles will also be assigned throughout the semester to provide examples of regression applications and discussions of best practices.

Recommended Texts

You will want to pick up a second regression text of your choosing that you can reference when you need more information or an alternate perspective. A few of the texts I like are listed below. You can often find older or international editions of these texts for significantly cheaper than the publisher's price. Many of these books are available online for free through the UT Library.

Basic Texts:

- Damodar N. Gujarati and Dawn C. Porter. *Basic Econometrics*. McGraw-Hill, 5th edition, 2008.

- A.H. Studenmund. *Using Econometrics: A Practical Guide*. Pearson, 7th edition edition, 2016.

Advanced Texts:

- Russell Davidson and James G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2003.
- William H. Greene. *Econometric Analysis*. Pearson, 8th edition, 2017.

Supplemental Texts:
(These are books that are less about the underlying math and more for the practice of econometrics or practical skills.)
Learning R:

- Florian Heiss. *Using R for Introductory Econometrics*. 2nd edition, 2020 (Available for free here: http://www.urfie.net)
- Hadley Wickham and Garrett Grolemund. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, 2017 (Available here: `https://r4ds.had.co.nz/i`

Theory and Practice:

- John Fox. *A Mathematical Primer for Social Statistics*. SAGE Publications, 2nd edition, 2020
- Will H. Moore and David A. Siegel. *A Mathematics Course for Political and Social Research*. Princeton University Press, 2013

Math refresh:

- Peter Kennedy. *A Guide to Econometrics*. Wiley-Blackwell, 6th edition, 2008.
- Gary King. *Unifying Political Methodology*. University of Michigan Press, 1989.

Writing:

- Leanne C. Powner. *Empirical Research and Writing: A Political Science Student's Practical Guide*. SAGE Publications, 2015

Administrative Issues

Academic Dishonesty

All violations of university academic conduct guidelines, including plagiarism, will be referred to the Dean of Students and will result in a zero on the assignment or exam in question. Plagiarism is the use of others' ideas or work without credit and/or presenting derivative work as one's own. This includes, but is not limited to, cutting and pasting from someone else's work or an internet source, failing to identify exact quotes, failing to cite a source for information that is only available from that source, failing to cite a source for an idea or argument you borrowed from that source, and turning in work that is not your own.

Disability Accommodations

Students with disabilities may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities (512-471-6259, http://www.utexas.edu/diversity/ddce/ssd/).

Religious Holidays

By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, I will give you an opportunity to complete the missed work within a reasonable time after the absence.

Course Outline

This is an approximate schedule that is subject to change depending on how quickly we move through the material.

January 9: Introduction

January 11: Linear Algebra

January 16: NO CLASS - MLK DAY

January 18: OLS in Matrix Form
- Wooldridge, Chapter 3.1-3.2 and Appendices D & E-1

January 23: Properties of OLS
- Gary King. Replication, replication. *PS: Political Science and Politics*, 28(3):444–452, 1995
- Gary King. Publication, publication. *PS: Political Science and Politics*, 39(1):119–125, 2006

January 25: Gauss-Markov Assumptions
- Wooldridge, Chapter 3.3-3.6 and Appendix E-2

January 30: Inference
- Wooldridge, Chapter 4.1-4.5 and Appendix E-3

February 1: Model Specification
- Problem Set #1 Due
- Wooldridge, Chapters 6.3
- Gary King. How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, 30(3):666–687, 1986

February 6: Interaction Terms

- Wooldridge, Chapters 6.2 and 7.4
- Thomas Brambor, William Roberts Clark, and Matt Golder. Understanding interaction models: Improving empirical analyses. *Political Analysis*, 14:63–82, 2006

February 8: Multicollinearity

- Kevin Arceneaux and Gregory A. Huber. What to do (and not do) with multicollinearity in state politics research. *State Politics and Policy Quarterly*, 7(1):81–101, 2007

February 13: OLS Diagnostics

- Sangit Chatterjee and Frederick Wiseman. Use of regression diagnostics in political science research. *American Journal of Political Science*, 27(3):601–613, 1983

February 15: Outliers and Influential Observations

- Problem Set #2 Due
- Wooldridge, Chapter 9.5
- Quan Li. Outlier, measurement, and the democracy-fdi controversy. *Quarterly Journal of Political Science*, 4:167–181, 2009

February 20: Heteroskedasticity

- Wooldridge, Chapter 8.1-8.3
- Gary King and Margaret E. Roberts. How robust standard errors expose methodological problems they do not fix, and what to do about it. *Political Analysis*, 23(2):159–179, 2015

February 22: Autocorrelation

- Wooldridge 12.1-12.5
- Luke Keele and Nathan J. Kelly. Dynamic models for dynamic theories: The ins and outs of lagged dependent variables. *Political Analysis*, 14(2):186–205, 2006

February 27: Weighted Least Squares

- Wooldridge, Chapter 8.4-8.5

March 1: Fixed Effects

- Problem Set #3 Due
- Wooldridge, Chapter 13.3-14.2
- Kosuke Imai and In Song Kim. On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 29(3):405–415, 2021

March 6: Catch up and Review

March 8: Midterm

March 13-17: Spring Break

March 20: Maximum Likelihood Estimation
- Ward and Ahlquist, Chapters 1-2

March 22: Generalized Linear Models
- Ward and Ahlquist, Chapters 5 & 7

March 27: Logistic Regression
- Ward and Ahlquist, Chapters 3 & 6

March 29: Count Models
- Problem Set #4 Due
- Ward and Ahlquist, Chapter 10

April 3: Categorical Models
- Ward and Ahlquist, Chapters 8-9

April 5: Introduction to Causal Inference
- Angrist and Pischke, Chapter 3
- Kosuke Imai, Luke Keele, Dustin Tingley, and Teppei Yamamoto. Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105(4):765–789, 2011

April 10: Differences-in-Differences
- Angrist and Pischke, Chapter 5
- Luke Keele. The statistics of causal inference: A view from political methodology. *Political Analysis*, 25:313–335, 2015

April 12: Regression Discontinuity
- Problem Set #5 Due
- Angrist and Pischke, Chapter 6

April 17: Instrumental Variables
- Angrist and Pischke, Chapter 4

April 19: Presentations

April 24: Presentations

Final papers due April 30