

GOVT 6029: Advanced Regression Analysis

Tues. & Thurs. 10:10–11:30am (Uris Hall G20)

Instructors

Peter Enns
peterenns@cornell.edu
740 Rhodes Hall

Office Hours: Tuesdays 1:15pm–3:15pm and by appointment

Sign up for office hours here (<https://enns.youcanbook.me/>) or email to make an appointment at a different time.

Teaching Assistant: Pauliina Patana (pauliina.patana@gmail.com)

Overview

GOVT 6029 is the second course in the Government Department's quantitative methods sequence. GOVT 6019 (Introduction to Probability and Applied Statistics) is a prerequisite for taking this course.

The course begins with model building and causal processes. These concepts will motivate the topics we cover throughout the rest of the course. We will spend most of the semester studying regression analysis using Ordinary Least Squares. We will focus on the assumptions of OLS, evaluating when these assumptions are violated, and how to proceed when such violations occur. We will also discuss challenges of regression based approaches for causal inference. We will end the semester with a brief introduction to Time Series Analysis and Maximum Likelihood Estimation (MLE) with limited dependent variables. Hopefully, the short time we spend on these two topics will be sufficient to provide the intuition behind time series and MLE and the knowledge to interpret research which uses these methodologies. The time we spend will not, however, be sufficient to write conference papers and articles which use time series or logit/probit. You should expect to take a specific course on the subjects and/or do extensive reading on the topics before using these methodologies.¹

The course assignments consist of homework to help you practice the concepts and skills we cover in class and a research paper where you apply the skills you learn to a topic of substantive interest. I expect the research paper to be suitable for presentation at a professional conference, such as the Midwest Political Science Association Conference.

¹To use these methods prematurely might be analogous to Luke leaving Dagobah in the *Empire Strikes Back*.

Objectives:

By the end of this course students will be able to:

- construct theoretical models that explain interesting political or social phenomena.
- use OLS regression to evaluate the predictions which stem from theoretical models.
- perform diagnostic tests to evaluate regression analyses.
- write a quantitative research paper suitable for a professional conference.

Texts

- *Basic Econometrics*. 2009. Damodar N. Gujarati and Dawn C. Porter. New York: McGraw-Hill

Statistical Software

Class examples will use Stata. Students should purchase Stata or use R.

Incomplete Policy

I only grant an incomplete if extenuating circumstances emerge (e.g., serious illness) and we consult about the situation during the course of the semester. Here is the relevant text from Cornell's incomplete policy (<http://courses.cornell.edu/content.php?catoid=12&navoid=2089>): "An incomplete (INC) signifies that a course was not completed before the end of the semester for reasons beyond the students control and acceptable to the instructor. Students must have substantial (normally at least 50 percent) equity in the course, be able to complete the remaining work, and have a passing grade for the completed portion... When a final grade is determined, it is recorded on the official transcript with an asterisk and a footnote explaining that this grade was formerly an incomplete."

Academic Integrity

Each student in this course is expected to abide by the Cornell University Code of Academic Integrity (<http://cuinfo.cornell.edu/aic.cfm>). Any work submitted by a student in this course for academic credit will be the student's own work.

Students agree that by taking this course all required papers may be subject to submission for textual similarity review to Turnitin.com for the detection of plagiarism. All submitted papers will be included as source documents in the Turnitin.com reference database solely for the purpose of detecting plagiarism of such papers. Use of Turnitin.com service is subject to the Usage Policy posted on the Turnitin.com site.

Evaluation

Your performance in this class will be assessed by your class participation, homework assignments, and a final research paper. Each component of your grade will be weighted as follows:

Class Participation	10%
Homework	50%
Research Paper	40%

Readings and Assignments

I have listed the assigned readings below. *Complete the day's reading before coming to class.* Most readings can be accessed through links on Blackboard. Otherwise, you can access articles through the Cornell Library's E-journal titles. If you are using the 4th edition of Gujarati, the corresponding page numbers are listed in footnotes. I may assign additional readings throughout the course.

- Week 1
 - Th. 1/26: Class Introduction

- Week 2
 - Tues. 1/31: Causal Processes and Statistical Modeling
 - * The Case for Working with Your Hands (<http://www.nytimes.com/2009/05/24/magazine/24labor-t.html>)
 - * Charles A. Lave and James G. March. 1975. *An Introduction to Models in the Social Sciences*, p.2-4, Ch.2, & p.60–61
 - * Bhatti and Erikson. 2011. “How Poorly are the Poor Represented in the U.S. Senate?” (p.224–226 (top))
 - * Willem Saris and Henk Stronkhorst. 1984. *Causal Modeling in Nonexperimental Research*, Ch.1 & Ch.2
 - * Morgan and Winship. *Counterfactuals and Causal Inference*. Ch. 1
 - Th. 2/2: Estimating Coefficients
 - * Gujarati & Porter, Ch.3 (esp. 3.2 through end) & p.97-101²

²p.107-112

- Week 3

- Tues. 2/7: Confidence Intervals and Hypothesis Tests

- * Gill, Jeff. 1999. “The Insignificance of Null Hypothesis Significance Testing.”
 - * Gelman and Loken. 2013. “The garden of forking paths.”
 - * Gelman, Andrew and Kaiser Fung. 2016. ‘The Power of the “Power Pose’: Amy Cuddys famous finding is the latest example of scientific overreach.” (http://www.slate.com/articles/health_and_science/science/2016/01/amy_cuddy_s_power_pose_research_is_the_latest_example_of_scientific_overreach.1.html).
 - * *Recommended:* Hoekstra, Morey, Rouder, and Wagenmakers. 2014. “Robust Misinterpretation of Confidence Intervals.”

- Th. 2/9: Measures of Fit, Estimation Theory

- * Gujarati & Porter Chs.7 and 8
 - * Achen, Christopher H. 1990. “What Does “Explained Variance” Explain?: Reply.” *Political Analysis*, 2(1): 173-184.

- Week 4

- Tues. 2/14: Interpreting Multiple Regression Coefficients

- * Anzia, Sarah F. and Terry M. Moe. 2015. “Public Sector Unions and the Costs of Government.” (focus on the analysis)
 - * Keele, Luke and Randolph T. Stevenson. 2014. “The Perils of the All Cause Model.”
 - * Ogorek, Ben. 2016. “How to Create Confounders with Regression: A Lesson from Causal Inference.” (<http://anythingbutrbitrary.blogspot.com/2016/01/how-to-create-confounders-with.html>).
 - * King, Gary. 2006. “Publication, Publication.”
 - * Gelman, Andrew. 2016. “Hey PPNAS... this one is the fish that got away - Statistical Modeling, Causal Inference, and Social Science.” (<http://andrewgelman.com/2016/09/18/another-item-for-uris-comment-section/>).

- Th. 2/16: Tests of Multiple Parameters

- * **Homework 1 Due**

- Week 5

- Tues. 2/21: NO CLASS (February Break)

- Th. 2/23: Adjusted R^2 and Categorical Predictors

- * **Paper Topic Due**

- Week 6

- Tues. 2/28: Introduction to Matrices
 - * Gujarati & Porter Appendix B (Note: Take enough time to “make sense” of these 12 pages)
- Thurs. 3/2: Matrix Computation and Applications to OLS
 - * **Data Proposal Due**
 - * James Stimson. “Professional Writing in Political Science: A Highly Opinionated Essay.”
 - * Michael Dunaway. 2017. “The Broken Hallelujah.” *The Folklore Project*. (<http://bittersoutherner.com/folklore-project/the-broken-hallelujah/#.WIAF6jDsNk0.twitter>).
 - * Jonathan Nagler. 1995. “Coding Style and Good Computing Practices.” *PS: Political Science and Politics*, 28(3): 488-492.

- Week 7

- Th. 3/9: Matrix Operations in Stata and Modeling Interactions
 - * Brambor et al. 2006. “Understanding Interaction Models: Improving Empirical Analyses.” *Political Analysis*, 14(1):63-82
 - * Breznau, Nate. 2015. “The Missing Main Effect of Welfare State Regimes: A Replication of Social Policy Responsiveness in Developed Democracies by Brooks and Manza.” *Sociological Science* (https://www.sociologicalscience.com/download/volume-2/august/SocSci_v2_420to441.pdf).
 - * Gujarati & Porter 288-290³
 - * *Recommended:* Berry, William D., Matt Golder, and Daniel Milton. “The Importance of Fully Testing Conditional Theories Positing Interaction”
 - * *Recommended:* Berry, William D., Jacqueline H.R. DeMerritt, and Justin Esarey. 2010. “Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?” *American Journal of Political Science*, 54(1):248-266.

- Week 8

- Tues. 3/14: Multicollinearity and Micronumerosity
 - * Gelman, Andrew and David Weakliem. “Of Beauty, Sex, and Power: Statistical Challenges in Estimating Small Effects.”
 - * Gujarati & Porter Ch.10

³p.310-312

- * *Recommended:* Lenth, Russell V. 2001. “Some Practical Guidelines for Effective Sample Size Determination.” *The American Statistician*.
- Th. 3/16: Dealing with Multicollinearity
 - * Arceneaux and Huber. 2007. “What to Do (and Not Do) with Multicollinearity in State Politics Research.” *State Politics and Policy Quarterly*, 7(1): 81-101.
 - * **Homework 2 Due**
- Week 9
 - Tues. 3/21: PCA, Ridge Estimation, Mixed Estimation
 - * Warwick, Paul V. 2002. “Toward a Common Dimensionality in West European Policy Spaces.” *Party Politics*, 8(1):101-122; (esp. 103-108).
 - * Breault, Kevin D. 1989. “A Reexamination of the Relationship between Religious Diversity and Religious Adherents” *American Sociological Review*, 54(6):1056-1059.
 - * Christopher Winship and Bruce Western. “Multicollinearity and Model Misspecification.” *Sociological Science*. (https://www.sociologicalscience.com/download/vol-3/july/SocSci_v3_627to649.pdf).
 - Th. 3/23: Heteroscedasticity
 - * Gujarati & Porter Ch.11
- Week 10
 - Tues. 3/28: Dealing with Heteroskedasticity
 - * **Homework 3 Due**
 - * King and Roberts. 2014. “How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It.”
 - * Hayes, Andrew F. and Li Cai. 2007. “Using Heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation.” *Behavior Research Methods*, 39(4):709-722.
 - * *Recommended:* Peter M. Aronow. 2016. “A Note on ‘How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It’ ” (<https://arxiv.org/pdf/1609.01774>).
 - Thurs. 3/30: Functional Form
 - * Gujarati & Porter Ch.13 & Ch.14
 - * Green, Donald Philip and Jonathan S. Krasno. 1988 “Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections” *American Journal of Political Science*, 32(4):884-907.

- Spring Break
 - Tues. 3/29: NO CLASS
 - Th. 3/31: NO CLASS

- Week 11
 - Tues. 4/11: Nonparametric Regression
 - * **Two Figures Due**
 - * Keele, Luke. 2008. *Semiparametric Regression for the Social Sciences*, John Wiley & Sons, Ltd. Chs. 1 and 2

- Week 12
 - Tues. 4/18: Instrumental Variables: Measurement Error and Causal Inference
 - * Sovey, Allison J. and Donald P. Green. 2011. “Instrumental Variables Estimation in Political Science: A Reader’s Guide” *American Journal of Political Science*, 55(1):188-200.
 - Th. 4/20: Influential Data Points
 - * Jennifer Wolak et. al. 2003. “California Dreaming: Replicating the ESA Model, Unusual Cases, and Comparative State Political Analysis.” *State Politics and Policy Quarterly*, 1(3): 255-272.

- Week 13
 - Tues. 4/25: Presenting Results
 - * 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):341-355.
 - * Kastlelec, Jonathan and Eduardo Leoni. 2007. “Using Graphs Instead of Tables in Political Science.”
 - Th. 4/27: Auto Correlation and Time Series Analysis
 - * **Homework 4 Due**
 - * Gujarati & Porter. Ch.12
 - * *Recommended:* Darmofal, David. 2006. “The Political Geography of Macro-Level Turnout in American Political Development.” *Political Geography*, 25(2):123-150.
 - * *Recommended:* Darmofal, David. “Spatial Econometrics and Political Science <http://people.cas.sc.edu/darmofal/SpatialEconometrics.pdf>

- Week 14

- Tues. 5/2: Time Series Analysis II

- * Gujarati & Porter Ch.17

- * Grant, Taylor and Matthew J. Lebo. 2016. “Error Correction Methods with Political Time Series.” *Political Analysis* 24:3-30.

- * Enns, Peter K., Nathan J. Kelly, Takaaki Masaki, and Patrick C. Wohlfarth. 2016. “Don’t Jettison the General Error Correction Model Just Yet: A Practical Guide to Avoiding Spurious Regression with the GECM.” *Research and Politics* Forthcoming.

- Th. 5/4: Maximum Likelihood Estimation and Interpreting Results

- * Gujarati & Porter Ch.15, p.102-106⁴

- * Grossback et al. 2005. “Comparing Competing Theories on the Causes of Mandate Perceptions” *American Journal of Political Science*, 49(2):406-419; esp. pgs. 415–417.

- * Maria Escobar–Lemmon “Political Support for Decentralization: An Analysis of the Columbian and Venezuelan Legislatures.” 2003. *American Journal of Political Science*, 47(4):683-697; esp. pgs. 690–694.

- Week 15

- Tues. 5/9: Interpreting Logit and Ordered Logit Regression Models

- Monday, May 15: **Research Paper Due**

⁴p.112–113 & 114–118