Outline of course on:

"Modern Regression Analysis"

Instructors: Dr. Guy D. Whitten, Laron K. Williams

Week 1

Overview

This course is designed to take students with a minimal background in statistics and mathematics and teach them the tools that they need to test their theories and produce presentations of their results for the top journals in Political Science, International Relations, Public Policy, and other related disciplines. Students are encouraged to bring their own data sets so that they can get hands-on experiences with applying the techniques covered in this course. Thus the emphasis is on making the transitions between theory, model speci_cation, and result presentation as seamless as possible.

The course is divided into three parts. The _rst part involves a thorough presentation of the logic and the central assumptions underlying the multiple ordinary least squares regression model. The second part focuses on issues that researchers typically encounter as they attempt to test their theories in a regression framework. The third part focuses on application and extension of the concepts covered in the _rst two parts. The topics covered in the third part of the course will depend on student interests and how much time we have.

This is a hands-on course, meaning that a major goal is to have students learn about techniques by putting them to work with statistical software. To facilitate this, we will have lectures on each topic followed by lab sessions. In these lab sessions students will have the option of working with their own data or working with data provided by the instructor. The main statistical software program that we will use for the labs is Stata. Guides to the free program R will also be available.

Class Schedule

We will spend as much time as necessary on each topic for this course. Topics 1 through 6 are essential topics that we will de_nitely cover. Additional topics will be covered based on available time and student interests.

Topic 1: Placing Quantitative Methods in a Research Program

- Course Introduction
- Working with Stata
- The Place of Methodology in a Research Program
- Some Rules of the Road For Conducting Statistical Analyses
- Taking Command of Mathematical Notation
- Getting to know your data

Readings:

Nagler, Jonathan (1995) “Coding Style and Good Computing Practices." The Political Methodologist, 6:2-8.

Chapters 1-5 of Kellstedt, Paul and Guy Whitten (2013) \The Fundamentals of Political Science Research, 2nd edition" Cambridge University Press

Topic 2: Two Variable Hypothesis Testing

- Some essential concepts from probability theory
- Covariance{Correlation

Readings:

Chapters 6-7 of Kellstedt, Paul and Guy Whitten (2013) \The Fundamentals of Political Science Research, 2nd edition" Cambridge University Press

Topic 3: Two Variable Regression

- Fitting a regression line
- Uncertainty in regression Part 1: Overall uncertainty -Goodness-of-Fit
- Introduction to regression simulation
- Uncertainty in regression Part 2: Uncertainty about individual components
- Assumptions and minimum mathematical requirements
- Introduction to regression simulation
- Inuential observations in regression
- Suppressing the intercept

Readings:

Chapters 8 & 12 of Kellstedt, Paul and Guy Whitten (2013) “The Fundamentals of Political Science Research, 2nd edition" Cambridge University Press

Topic 4: X -Y controlling for Z

- All the di_erential calculus you need for right now
- Extending to multiple regression
- Introduction to matrix algebra
- OLS in matrix algebra
- F-distribution and F-tests
- Presenting model results
- Omitted variable bias
- OLS is “BLUE"

Readings:

Fair, Ray C. (1988) “The e_ect of economic events on votes for president: 1984 update" Political Behavior, 10:168-177.

Kellstedt and Whitten Chapter 9

Jann, Ben (2014) “Plotting regression coe_cients and other estimates" Stata Journal,

14:708-737.

Topic 5: Relaxing and Testing Regression Assumptions I

- Multicollinearity and Micronumerosity
- Introduction to Clarify
- Being smart with dummy variables
- Inuential Cases in Multiple Regression
- Functional form adjustment
- To standardize or not

Readings:

Kellstedt and Whitten Chapter 10

King, Gary, Michael Tomz, and Jason Wittenberg (2000) \Making the Most of Statistical Analyses: Improving Interpretation and Presentation." American Journal of Political Science, 44:347-61.

Tomz, Michael, Jason Wittenberg, and Gary King (2003)”Clarify: Software for Interpreting and Presenting Statistical Results."

Topic 6: Relaxing and Testing Regression Assumptions II

- Heteroscedasticity
- “Curing" Heteroscedasticity with FGLS
- Interactive models

Readings:

Chapter 2 of Kam, Cindy and Robert Franzese (2009) Modeling and Interpreting Inter-active Hypotheses in Regression Analysis. University of Michigan Press.

Brambor, Thomas, William Clark and Matt Golder (2006) \Understanding Interaction Models: Improving Empirical Analyses." Political Analysis, 14:63-82.

Berry, William D., Matt Golder, and Daniel Milton (2012) \Improving tests of theories positing interaction." Journal of Politics 74.3: 653-671.

Topic 7: Time Series Analysis

- Introduction to time series data and notation
- Introduction to using Stata to analyze time series data
- Non-stationarity and unit roots
- Overview of the threats to inference in time series analysis
- Traditional treatments of autocorrelation in regression models
- Interpreting time series regression models
- To lag or not to lag

Reading:

Kellstedt and Whitten pp. 256-269.

Beck, Nathaniel (1991) \Comparing Dynamic Speci_cations: The Case of Presidential Approval," Political Analysis III.

Keele, Luke and Nathan Kelly (2006) \Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." Political Analysis, 14:186-205.

De Boef, Suzanna and Luke Keele (2008) \Taking Time Seriously." American Journal of Political Science, 52:184-200.

Topic 8: Pooled Time Series Analysis

Note: This topic can not be covered without first having covered \Time Series Analysis."

- To pool or not to pool?
- Fixes with fixed effects?
- Panacea Corrected Standard Errors?
- But wait, there's more

Assigned Reading:

Stimson, James (1985) “Regression in Time and Space: A Statistical Essay" American Journal of Political Science, 29:914-947.

Beck, Nathaniel and Jonathan Katz (1995) \What To Do (and Not To Do) with Time Series Cross-Section Data." American Political Science Review, 89:634-47

Williams, Laron K. and Guy D. Whitten (2012) \But Wait, There's More! Maximizing Substantive Inferences from TSCS Models." Journal of Politics, 74:685 693.

Topic 9: Qualitative Response Models

Readings:

Kellstedt and Whitten pp. 247-255.

Whitten, Guy D. and Harvey D. Palmer (1996) \Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands." American Journal of Political Science, 40:231-60.

Topic 10: Spatial Regression Models

Readings:

Neumayer, Eric and Thomas Pl•umper (2010) \Making spatial analysis operational: Commands for generating spatial-e_ect variables in monadic and dyadic data" The Stata Journal 10.4:1-21.