

# WSSR

## WORKSHOPS ON SOCIAL SCIENCE RESEARCH

### Modern Regression using OLS

**Dr. Guy Whitten**

Professor and Cornerstone Fellow,

Texas A&M University

June 1-2, 2017

9:00am – 4:30pm

#### Workshop Outline

This course is designed to take students with a minimal background in statistics and mathematics and teach them the tools that they need to begin 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 specification, and result presentation as seamless as possible.

The course is divided into two parts. The first 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.

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.

#### **Topic I - Placing Quantitative Methods in a Research Program**

##### 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.

<http://0-www.jstor.org/mercury.concordia.ca/stable/420315>

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
- Influential observations in regression
- Introduction to regression simulation
- 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 - $X \rightarrow Y$ controlling for Z

- All the differential calculus you need for right now
- Extending to multiple regression
- Some basics of matrix algebra
- OLS in matrix algebra
- F-distribution and F-tests
- Presenting model results
- Omitted variable bias

### Readings:

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

Available on course reserves: <https://reserves.concordia.ca/ares/>

Kellstedt and Whitten Chapter 9



## Topic 5 - Relaxing and Testing Regression Assumptions

- Multicollinearity and Micronumerosity
- Introduction to Clarify
- Being smart with dummy variables
- Influential Cases in Multiple Regression
- Functional form adjustment
- To standardize or not
- Heteroscedasticity
- “Curing” Heteroscedasticity with FGLS
- Interactive models

### 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. <http://0-www.jstor.org.mercury.concordia.ca/stable/2669316>

- 
- 
- Tomz, Michael, Jason Wittenberg, and Gary King (2003) "Clarify: Software for Interpreting and Presenting Statistical Results." <https://gking.harvard.edu/files/clarify.pdf>
- Chapter 2 of Kam, Cindy and Robert Franzese (2009) *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. University of Michigan Press.  
<http://reserves.concordia.ca/ares/ares.dll/plink?14EEF89C>
- Brambor, Thomas, William Clark and Matt Golder (2006) "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis*, 14:63-82.  
<http://0-www.jstor.org/mercury.concordia.ca/stable/25791835>
- Berry, William D., Matt Golder, and Daniel Milton (2012) "Improving tests of theories positing interaction." *Journal of Politics* 74.3: 653-671.  
<http://0-search.ebscohost.com/mercury.concordia.ca/login.aspx?direct=true&db=poh&AN=78166626&site=ehost-live&scope=site>