# Pol 552: Applied Political Research
# Spring 2015

**Basic Information**

Class meets: Wednesday, 1:00-3:30 p.m.
Instructor: Dr. Heather Ondercin
Office: 233 Deupree Hall
E-mail: ondercin@olemiss.edu
Office hours: Tuesdays 1:00-2:30, Thursdays 9:00-10:30, and by appointment.

**Course Description**

This is the second class in a two-course graduate level sequence on quantitative methodology for political science. This course provides an overview of several common methods for multivariate data analysis. It will also help you recognize which statistical methods are appropriate for different types of data generating processes and data structures. The primary focus of this class is the linear model estimated via ordinary least squares (OLS) and its extensions. We will cover the assumptions associated with OLS, how to test these assumptions, and how to proceed when the assumptions are violated. We will also cover the basics of model specification. We will conclude class with a brief overview of models for binary and ordinal dependent variables. This class is designed to provide students with a foundation in the statistics underlying these methods of estimation, but also an intuition and ability to apply these methods in their own research.

This class is designed to be a class in applied methods. The best way to learn how to conduct multivariate analysis is to actually conduct analysis. Class will be divided between lecture/discussion and lab practicum. During the lab practicum students will work in STATA, a statistical analysis program, to analyze data and apply the material covered in lecture. In addition to the lab practice, students will gain additional experience working with these methods by completing a series of homework assignments and a final research paper.

**Learning Objectives**

If you put in the effort this course demands (which is a lot), then you can expect to master many statistical concepts and tools for analyzing quantitate data.

- Identify the correct model for the type of dependent variable and data structure.
- Estimate linear (and a few non-linear) regression models using STATA.
- Interpret estimates from linear (and a few non-linear) regression models.
- Diagnose and, when possible, correct problems for estimation and inferences in these models.

**Class Expectations**
Attendance is mandatory for every class meeting. If you are absent, you will not receive credit for participation on that day. Additionally, we cover a lot of material in a single class. I do not make my notes or slides available to students; if you miss a class for some reason it is your responsibility to make arrangements with other students in the class.

I expect that students will have completed the assigned readings before coming to class. Reading for a stats class is considerably different than reading for a substantive graduate seminar. There will be considerable fewer pages of reading every week, but it will take you longer to read a page of in your stats book than a page in a substantive book. I strongly recommend taking notes on the readings. Each student should be prepared to ask and answer questions on the course material in class.

**Class Requirements**
Grades (using a plus/minus letter grade system) will be based on:

- 5 Problem Sets (50%)
- Application Readings (20%)
- Replication and extension project (20%)
- Participation (10%)

**Problem Sets** There are will be five problems sets that you will need to complete throughout the semester. Each problem set is worth 10% of your final grade. These assignments will focus on data analysis and interpretation. The lab practicums will prepare you for the work that you will complete in these assignments.

Warning: Statistical homework is fundamentally different than homework that you may have done for other classes. Most importantly, you should not leave it until the last minute! You will find that some aspects of the homework can be resolved quickly with a little help (e.g., some computer glitch), but these problems can take your hours to resolve by yourself.

A completed homework assignment will consist of two work products: 1) Your write up; and 2) your STATA .do file. Each of you will have share a dropbox folder with me. You will upload the .do file and write-up to the dropbox folder by the due date for the assignment. No late assignments will be accepted.

The write-up should be completed in a word processor (i.e. Microsoft Word or LaTex). In the text of the write-up you should completely answer the questions asked in the assignments. Your answers should be complete sentences, using proper grammar and spelling. Your answers should demonstrate a mastery of concepts. Make sure to clearly flesh out all of your answers and discuss in detail the tests and any interpretation. When figures

and tables are required, they should be nicely formatted and clearly labeled. Pick up any political science journal to get an idea of how your tables and figures should look. Do not turn in STATA output that has be directly copied and pasted into your document for the tables. Because the write-up must be word processed you should familiarize yourself with the means to insert equations, mathematical functions, and other symbols (e.g., Greek letters) in your documents. MS Word users will find this easy enough with that program's built-in equation editor and options for symbols. All math work must be nicely formatted (i.e., one equation per line; using appropriate symbols; on a separate line from the text and centered) and complete (i.e., show all your work).

Your stata .do file should be clear of errors. I should be able to change the file pathways and run the .do file to reproduce your results.

I encourage students to work together on the assignments, but you should write your own solutions (this includes computer work), and I strongly suggest that you make a solo effort at all the problems before consulting others. If you work in groups, you should include the names of your coworkers in your write-up. You will be expected to follow all university policies regarding academic honesty. If you do not follow these guidelines, your problem set will receive a zero.

**Application Readings**
You will complete four application readings, write a short paper on these readings, and be prepared to discuss the reading in class. Each application reading is worth 5% of your final grade. You will select one of the readings from the list of readings provided. After completing the reading, you will write a short paper between 2-3 pages in length. The paper should (1) summarize the argument and findings in the article, (2) identify and discuss the data structure, (3) identify and discuss the methods selected. The summary of the argument and findings should take up no more than 1/3 of the paper. The discussion of the methods should draw connection to the substantive material being discussed in class and in the assigned readings. We will be discussing these papers in class. For your application reading you are expected to present and discuss the paper with your classmates. The papers associated with these readings are due on the day they are listed on the syllabus.

**Final Research Project**
The final project will require you to apply the techniques learned in this course to an existing data set or to a set of data you assembled. For this project you have two options: (1) a replication and extension of published research or (2) an original research paper. The topic should be chosen in consultation with me. Paper proposals will be due on February $25^{th}$. The final paper for either options should range between 10-15 pages, not including tables, graphs, and citations. Both papers should resemble a journal article minus the literature review and on methods steroids. The final paper should include a brief introduction, theoretically grounded expectations, research design section that discusses the data and methods

used to conduct the analysis, results section, and conclusion. In addition to the paper, you will submit both the data used for the project and all .do files needed to replicate the material in the paper. Final papers are due by 9am on May $7^{th}$. Students should turn a hard copy of the paper into my mailbox and upload all other material to our shared dropbox folder by that time. Late papers will not be accepted. You cannot turn in a paper for this class that you turned in for another class.

Option 1: Replication and extension of existing research.
This project requires you to replicate and extend the data analysis of a published journal article. The replication should produce the same results as those in the published article. The extension of the data analysis can be one of two forms. One type of extension is to cross-validate the data analysis in a larger or alternative sample. Another type of extension is to cross-validate the data analysis based on an alternative model specification or estimator. The purpose is to assess robustness of previous findings in the presence of confounding variables or alternative estimators. The analysis should be motivated theoretically.

Option 2: Original data analysis.
This project will require you to identify an original research question and test that question using the techniques covered in this class. The analysis should be theoretically motivated. This paper is an expanded version of the research design and results section in a journal length manuscript. This paper may overlap with a paper you are completing for another class, but you should not turn in the same paper to both classes. Rather, this paper needs to be specific to the requirements of this class. Additionally, you will need to obtain permission from your other instructor that you are using the some of the same work for two classes.

**Participation**
You will not get much out of this class if you are a passive observer. You should come to class having read the material and prepared to discuss the different topics. You should always be thinking about how this relates to your research questions and interests. In addition to being an active participant in class, part of the participation grade will be based on presentations of applied readings. You are expected to be an expert on the article you selected for discussion that day. I will choose someone for each article to start off the discussion. Just because you are not chosen, you are not off the hook. You will be expected to ask questions and contribute to the conversation.

## Books and Materials

There are two required books for the class.

- *Introduction to Econometrics, Fourth Edition.* by Christopher Doughtery. Oxford University Press. ISBN: 978-0199567089

- *Statistics with Stata: Version 12, Eighth Edition* by Lawrence C. Hamilton. Cengage. ISBN: 978-0-8400-6463-9 (The STATA bookstore has the cheapest price: `http://www.stata.com/bookstore/statistics-with-stata/`)
- A copy of STATA on a laptop that you can bring to class. The department will be purchasing you a copy of STATA.

I highly recommend that you purchase both of these books versus renting the books. It is helpful to develop a statistics and methods reference library. You will find this a valuable resource when working on papers for other classes, your master's or practicum papers, your dissertation, and future research. Other books that you might want to consider adding to your methods library in the future are:

- Good Econometrics Books (Mid-range in terms of math and technical details)
  - *A Guide to Econometrics, Sixth Edition* by Peter Kennedy. Wiley-Blackwell.
  - *Basic Econometrics, Fifth Edition* by Damodar Gujarati and Dawn Porter. McGraw-Hill/Irwin.
  - *Introductory Econometrics: A Modern Approach, Fifth Edition* by Jeffrey M Wooldridge. Cengage Learning. (This is probably my favorite. If you purchase a single additional book I would highly recommend this one.)
- Additional Good Econometrics Texts that are more technical.
  - *Microeconometrics: Methods and Applications* by Colin Cameron and Pravin K. Trivedi. Cambridge University Press. (For a more technical book this also is nicely intuitive. There is also a STATA companion for this book that is really nice).
  - *Econometric Analysis, Seventh Edition* by William H. Greene. Prentice Hall. (The most technical).
- Good guides or other resources
  - Achen, Christopher H. 1982. *Interpreting and Using Regression* Sage University Press. (This is one is a series of "little green books" from sage on methods. They are short, usually non-technical, and really helpful).
  - Institute for digital research and education at UCLA has an excellent information about STATA: `http://www.ats.ucla.edu/stat/stata/`
- Beyond OLS
  - Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Sage. (There is also a STATA book that is designed as a companion by Long and Freese).

– King, Gary. 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. University of Michigan Press.

## Tentative Course Schedule[1]

Week 1 (January $21^{st}$): Introduction and Review

- Christopher H. Achen. "Advice for Students Taking a First Political Science Graduate Course in Statistical Methods." *TPM*. (I know this is not your first stats class, but it is still a good read).
- Nagler, Jonathan. 1995. "Coding Style and Good Computing Practices." *The Political Methodologist*. 6(2):2-9. This is available in pdf format online at http://polmeth.wustl.edu/. Select The Political Methodologist and the correct volume.

Week 2 (January $28^{th}$): Introduction to Regression

- Dougherty, Sections R1-R15.
- Skim Hamilton Chapter 1 & 2

Week 3 (February $4^{th}$): OLS, Gauss-Markov Assumptions & Hypothesis Testing

- Dougherty Chapter 1 & 2
- Skim Hamilton Chapters 3 & 5
- Application Reading on Bivariate Regression

Week 4 (February $11^{th}$): Multivariate Regression

- Dougherty Chapter 3
- Start reading Hamilton Chapter 7

Week 5 (February $18^{th}$) Working with STATA

- Problem Set 1 Due.
- Hamilton Chapter 7, pages 163-181.

Week 6 (February $25^{th}$):Model Specification I: Outliers, Influential Data, and Multicollinearity.

- Hamilton, Chapter 7 pages 190-202
- Application reading on multivariate regression.

[1]Please pay attention to e-mail and announcements in class regarding any schedule changes.

- Paper Proposal Due.

Week 7 (March $4^{th}$): Model Specification 2: Linearity and Omitted Variable Bias.

- Dougherty Chapter 4
- Dougherty Chapter 6

Week 8 (March $11^{th}$): Spring Break

Week 9 (March $18^{th}$): Model Specification 3: Dummies and Interaction Terms.

- Dougherty Chapter 5
- Hamilton Chapter 7, pages 181-190.
- Brambor, Thomas, Wiliam Robers Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analysis" *Political Analysis* 14: 63-82.
- Problem Set 2 Due.

Week 10 (March $25^{th}$) Post-Estimation and Interpretation on Regression

- Application reading on dummy variables or interaction terms.
- King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation" *American Journal of Political Science* 44: 341355.
- Clarify documentation (`http://gking.harvard.edu/clarify`)
- Look over stata code and material on: `https://files.nyu.edu/mrg217/public/interaction.html`
- Hamilton Chapter 14

Week 11 (April $1^{st}$): Problems in the Errors 1: Heteroskedasticity

- Dougherty Chapter 7
- Hamilton Chapter 7, pages 202-215
- Problem Set 3 Due.

Week 12 (April $8^{th}$): Problems in the Errors 2: Serial Correlation

- Dougherty Chapter 12
- Hamilton Chapter 12, pages 351-371

Week 13 (April $15^{th}$): Work day

- Problem Set 4 Due.

Week 14 (April $22^{nd}$) Models for Binary Dependent Variables

- Long Chapter 3 & 4
- Hamilton Chapter 9, pages 251-268
- Application reading on binary dependent variables.

Week 15 (April $29^{th}$) Models for Ordinal Dependent Variables

- Long Chapters 5
- Hamilton Chapter 9, pages 269-282
- Problem Set 5 Due.