

## Data Analysis II – (POLS 501)

Spring 2016

(This syllabus is current as of January 14, 2016, but is subject to change at any time.)

Instructor: Dr. Noah Kaplan

Class: Thursday, 6 – 8:30 pm, BSB 1115

Office: 1122B BSB, ext. 6-5156

E-mail: njkaplan@uic.edu

Office Hours: Tuesdays 9:30 – 11:30 am (and by appointment).

TA: Semih Patan

Office Hours: Wednesdays 1 – 3 pm in BSB 1145 (and by appointment).

E-mail: spatan3@uic.edu

This course will review the principles and basic methods of multivariate analysis used in political science and public policy research. Its purpose is to intensively examine the concept of multivariate regression and provide an understanding of the various issues and problems which can occur in the use of multiple regression analysis and approaches to addressing such issues/problems.

This course assumes familiarity with the basic concepts of levels of measurement, probability, hypothesis testing, and multiple (OLS) regression. You should be familiar with a statistical software package such as SPSS, SAS or Stata. This course aims at familiarizing you with the underlying assumptions of regression for data analysis so that you will be able to evaluate and undertake quantitative research.

**Class and Lab Sessions:** The class meets once a week on Thursday. Class will regularly meet in the computer lab BSB 4133. I will announce beforehand if we are to meet the following class in BSB 4133. Students *must* attend no matter where we meet (i.e., whether we meet in BSB 1115 or BSB 4133, attendance is required).

### Readings and Assignments:

Readings for each section of the course are specified in the attached syllabus. The books for this course are:

#### Required Texts

- Damodar Gujarati, *Basic Econometrics* (3<sup>rd</sup> ed. or 4<sup>th</sup> ed.)
- J. Scott Long & Jeremy Freese, *Regression Models for Categorical Dependent Variables Using Stata 3<sup>rd</sup> Edition* (Stata Press; check price at Stata's bookstore at: <http://www.stata.com/bookstore/stata-press-books/>)
- Fred Pampel, *Logistic Regression: A Primer* (#132)

Recommended Texts -- All the recommended books are from the series "Quantitative Applications in the Social Sciences" by Sage Publishers. I have not ordered these books. If you wish to purchase these books, you can order them from any bookseller (real or virtual) or you may order them directly from Sage. The URL for Sage is <http://www.sagepub.com>.

- Michael Lewis-Beck, *Applied Regression* (#22)
- Christopher Achen, *Interpreting and Using Regression* (#29)
- J. Jaccard and R. Turrisi, *Interaction Effects in Multiple Regression* (#72)
- J. Aldrich and F. Nelson, *Linear Probability, Logit and Probit Models* (#45)
- James Davis, *The Logic of Causal Order* (#55)
- C. Ostrom, *Time Series Analysis* (#9)
- J. Scott Long, *Regression Models for Categorical and Limited Dependent Variables* (Sage Publications)

Optional/Additional Texts of Potential Use

- Peter Kennedy, *A Guide to Econometrics* (paperback)
- A. H. Studenmund, *Using Econometrics: A Practical Guide*

If you are finding Gujarati difficult to follow, you may wish to consult Kennedy or Studenmund for another perspective/voice on the material.

The course requires the use of computers for most assignments. We will be using a statistical program called Stata.

**Hardware and software:** We use Stata for Windows extensively in this class; this program is loaded on the computers in department's graduate student computer lab, as well as many other computers around the university. You may wish to buy Stata (you can purchase perpetual license of Stata/IC 14 for \$198, or for Stata/SE 14 for \$395 (or you can purchase a six-month license or a one year license of State/SE 14; for further information, see <http://www.stata.com/order/new/edu/gradplan.html>). To determine which version of Stata best suites your needs, see: <http://www.stata.com/products/which-stata-is-right-for-me/>. You do **not** need to own the software; it just gives you the convenience of working at home and/or on a laptop. Stata is in the political science graduate student lab and in the computer lab we will be using during the semester (BSB 4133)

**Teaching Assistant:** One Teaching Assistant is associated with this course. The TA will hold weekly office hours and will be available by appointment on an as need basis.

**Grading** is based on the 5 required papers as follows:

- Paper #1 – OLS analysis of a multivariate model (15%)  
Due Feb. 4<sup>th</sup>
- Paper #2 – Re-estimate the multivariate model testing and correcting for heteroskedasticity (20%)  
Due Feb. 18<sup>th</sup>
- Paper #3 – Dichotomous Dependent Variable Analysis (20%)  
Due March 10<sup>th</sup>
- Paper #4 – Ordinal Analysis (20%)  
Due March 31<sup>st</sup>
- Paper #5 – Multinomial Analysis (20%)  
Due April 14<sup>th</sup>
- Class Attendance (5%)

**Note: NO late papers will be accepted.**

**Lab Assignments:** There will be lab assignments the first couple of weeks of the semester. These assignments are intended to help refresh your memory regarding Stata and OLS. You must hand-in the assignments to the TA in lab the following week. Lab assignments will be posted on the course web page.

**Note 1:** Weekly assignments **MUST** be done by each individual. Assignments must be turned in at the beginning of class on the day that they are due. **NO late work will be accepted.** Graded problems will be returned in lab.

Attendance is taken at the beginning of each class period. If you miss more than two classes, your course grade will be reduced by one letter grade for each additional class you miss.

**Academic Etiquette:** Do not carry on side conversations or read the newspaper during class. Doing so is disrespectful to me and your classmates, and I will ask you to leave the room. **Turn OFF all cell phones before class begins.**

**Cheating and Plagiarism:** I do not tolerate plagiarizing or cheating of any kind. Such behavior will result in **failing** the course and other disciplinary action. Cheating and plagiarism in this class will be punished to the maximum extent possible. Any violation may result in expulsion from the University. All students are expected to observe the University of Illinois at Chicago's rules against cheating and plagiarism. See the section on "Academic Integrity/Academic Dishonesty" in the University of Illinois at Chicago's Undergraduate Student Academic Policy Guide 2009-2011 (<http://www.uic.edu/depts/oa/SMAAPP/guide.pdf>) for a full statement regarding UIC's rules against cheating and plagiarism. The relevant regulations and procedures can be

found at <http://www.uic.edu/depts/dos/docs/Student%20Disciplinary%20Policy.pdf>. The department of political science provides a full definition of plagiarism and guidelines on how to avoid plagiarism at: <http://www.uic.edu/depts/pols/Plagiarism1.pdf>. Cheating and plagiarism in this class will be punished to the maximum extent possible.

**ADA Statement:** The American with Disabilities Act (ADA) is a federal antidiscrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please see <http://disabilityresources.uic.edu/> and contact the UIC's Disability Resource Center (for contact information and more, see <http://drc.uic.edu>). Then follow-up with me as appropriate.

## **Data Analysis II – Political Science 501**

### **COURSE OUTLINE**

Section 1: January 14<sup>th</sup> through February 11<sup>th</sup> (first 5 weeks, inclusive)

#### **OLS Regression**

Gujarati chapters 2-8, 10, 11

Michael Lewis-Beck, *Applied Regression* (#22)

Christopher Achen, *Interpreting and Using Regression* (#29)

#### **Paper #1: Due Thursday, Feb. 4<sup>th</sup>**

Estimate a model with one or two key independent variables (and the standard seven variables controlling for demographic, socio-economic and political factors), using the data available in the lab, or *using your own data*. Discuss your hypothesis concerning the relationships of interest to you. Report the relevant statistics, and discuss your findings. Test for an interaction, and analyze outliers. Transform variables, if necessary. Include a 200 word abstract. The paper should be approximately 10 pages in length.

#### **Paper #2: Due Thursday, Feb. 18<sup>th</sup>**

Reestimate the multivariate model, analyze residuals. Test and correct for heteroskedasticity. Compare results with the original model. Include an abstract. The paper should be approximately 10-12 pages in length.

Section 2: Feb. 18<sup>th</sup> through March 3<sup>rd</sup> (3 weeks, inclusive)

#### **Models with Dichotomous Dependent Variables**

Pampel, entire (also Recommend Aldrich and Nelson)

Gujarati, chapters 14 & 15

Long & Freese, chapters 5 & 6

**Paper #3: Due Thursday, March 10<sup>th</sup>**

Estimate a model for a dichotomous dependent variable using OLS and one of the other methods discussed in the class. Explain why the method you choose is superior to others (e.g., logit is superior to OLS given your dependent variable). Discuss and compare results. Note any substantive discrepancies. Include an abstract. The paper should be approximately 8-10 pages in length.

Section 3: March 10<sup>th</sup> – March 17<sup>th</sup> (two weeks, inclusive)

**Models with Ordered Dependent Variables**

Long & Freese, chapter 7

**Paper #4: Due Thursday, March 31<sup>st</sup>**

Estimate an ordinal model using your own data (i.e., 2012 NES dataset). Discuss/interpret results. Include an abstract. The paper should be 8-10 pages in length.

Spring Break, Thursday, March 24<sup>th</sup>.

Section 4: March 31<sup>st</sup> – April 7<sup>th</sup> (two weeks, inclusive)

**Models with Nominal Outcomes**

Long & Freese, chapters 8

**Paper #5: Due Thursday, April 14<sup>th</sup>**

Estimate a multinomial model using your own data (i.e., 2012 NES dataset). Discuss/interpret results. Include an abstract. The paper should be 8-10 pages in length.

Section 5: April 14<sup>th</sup> – April 28<sup>th</sup> (three weeks, inclusive)

**Brief Introduction to Autocorrelation & Time Series Using OLS**

Gujarati, chapters 12, 17

C. Ostrom, *Time Series Analysis* (#9)

**No paper.** There will be a couple of assignments (due the 21<sup>st</sup> and 28<sup>th</sup> of April, respectively).

THE LAST DAY OF CLASS IS THURSDAY, April 28<sup>th</sup>.