

PLCS 502: Statistical Methods for Political Research

Fall 2005

Monday (lab) 1:00-2:30 124 Pond

Wednesday 1:00-2:30 124 Pond

Friday, 9:05-11:00 122 Pond.

Suzanna De Boef

Phone: 863-9402

Email: sdeboef@psu.edu

Office: 320 Pond

Office hours: Wednesday 9:00-11:00, Friday 11:00-12:00

webpage: <http://polisci.la.psu.edu/faculty/deboef/welcome.htm>

TA: Elena Kirtcheva

Phone: 777-5839

esk131@psu.edu

Office: Pond 214

Office Hours: Tuesday and Thursday 1:00-2:30

Course Description

This course is part of the methods sequence that is required of doctoral students in political science. It is designed with two key objectives in mind. First, all students need to understand the statistical methods widely applied in political science research. Second, students need to be able to apply appropriate statistical methods to analyze substantive problems in political science. The material we will cover in this course and that which will follow in 503 and beyond will set you on your way in this task.

We will focus on how to use statistical methods to estimate and test causal relationships. I have divided the course into three parts. We begin with essential topics in probability theory as well as statistical inference and hypothesis testing. You will need to (re)familiarize yourself quickly with basic concepts of central tendency, variability, random samples, correlations, etc. From here we cover how to use statistics to evaluate hypothesized causal relationships among variables through simple and multiple regression analysis. In this section you will learn how to estimate parameters and test hypotheses about them in bivariate and multivariate regression, understand the statistical properties of the OLS estimator under the set of Gauss-Markov assumptions, and understand the statistical and intuitive meaning and consequences of violations of these assumptions. You will also learn a variety of ways to diagnose and correct for violations of the assumptions so that regression remains an appropriate tool for inference in practice. In the third part of the course we will cover extensions of simple regression analysis in terms of assumption violations. These will include dealing with limited dependent variables, time series data, and endogeneity, for example.

Course Requirements and Evaluation

You will have succeeded in this course if you understand the material we cover both algebraically and intuitively, that is if you can present the statistical idea using both algebraic notation and numeric values and also explain the ideas represented by the notation. These two skills are complementary and fundamental to appropriate applications of quantitative methods and your success in future methods courses as well.

The second mark of success is the ability to use the statistical ideas to interpret and evaluate quantitative studies in political science journals and also to conduct a quantitative research project independently. The former requires that you understand the material covered in the course, the latter requires proficiency in data management and statistical analysis using a statistical software package. For this course that package will be Stata (with the occasional smattering of R). Homework assignments and exams are designed to train and evaluate you in these two aspects.

Elena Kirtcheva, who will be serving as the TA for the course, and I have selected problem sets, designed exams, a replication research project, and a semester long project for the class to promote this type of understanding.

Students are expected to finish the required readings before coming to class. Each student should be prepared to ask and answer questions on course materials in class. Performance in class discussion and questions about course materials will be graded at the end of the semester. Class participation is also highly correlated with student performance on other components of the course requirements. It is more interesting and efficient to participate actively in a graduate course than simply to get by. A student who misses classes regularly is unlikely to perform well in the course.

Weekly problem sets will be distributed throughout the semester. They require solving mathematical problems, computer exercises, diagnosis and correction of regression assumption violations and applications in advanced topics. In these exercises, it is important to carry out multiple tests (and when possible, go beyond the scope of the lecture) and to discuss carefully the implications of the tests for statistical inference and substantive interpretation. The problem sets are due in class each Monday, as specified in each weekly assignment. **No late assignment will be accepted.** Taking into account unavoidable scheduling conflicts with other classes, I will allow each student to skip one and only one homework assignment.

You will be creating a class reference book, a sort of encyclopedia to 502. Over the course of the semester each student will be responsible for writing and presenting three entries. Entries will have 3 parts. The first will be definitional and require both notational and conceptual exposition. The second part will require you to develop the definition, must be entirely composed in words of your own, and should convey a level of understanding beyond that apparent from the definition alone. You should consult 3-4 textbooks before closing them all and writing your own words on the topic. The third part should give an example or two in which the term would be of use in political science. Entries will be presented at the beginning of the class, one each Monday and Wednesday, two on Fridays beginning on the 7th of September. Presentations will be about 10 minutes long. Materials that serve as the basis of the presentation must be made available to your classmates and may contain more information than is presented. Questions about the entries, which will be tied directly to the regular course material, will be part of each exam.

A research project based on a substantive problem in your major field of study is required. The project requires you to **replicate and extend** the data analysis in 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 example. The sample may be expanded in temporal or cross sectional units, motivated by some theoretical intuition. 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 also be motivated theoretically. More details of the project will be discussed in class after the first exam. The project should have about 10 pages of text. The topic must be chosen in consultation with me.

Three in-class exams cover the three parts of the course materials. Content of the exam will be based on the textbook, course lectures, the problem sets, and the class reference book entries.

Other than the regular lectures, the course also has a required lab session each week. Elena Kirtcheva will lead the (Monday) lab session. During each lab session, Elena will discuss problems and issues in the weekly homework assignment, lecture on programming in Stata (and occasionally R), and help with students' practice with Stata during each session.

Grading

Problem sets: 30%

Reference Entry & Presentation: 5% each $\times 3 = 15\%$

Research Project: 20%

Exams: 10% each $\times 3 = 30\%$

Class Participation: 5%

Books, etc.

- David Howell Fundamental Statistics for the Behavioral Sciences. 5th edition. 2004 Thompson.
- Christopher Dougherty. Introduction to Econometrics. 2th edition. Oxford University Press.
- Lawrence Hamilton. Statistics with Stata, updated for version 8 (2004 edition) Duxbury Press.

Course Calendar

Part I Baby Stats

Week 1 August 31-September 9: Introduction, *really* basic concepts, graphic display, and measures of central tendency

- No class the 5th, Labor Day holiday
- Read Howell Chapters 1-4.
- Homework Assignment One Due Monday, September 12 at the start of class.

Week 2 September 14-19: Measures of Variability, Basic Probability Stuff.

- Read Howell Chapters 5 and 7
- Review from Dougherty, Expected Values
- Homework Assignment Two Due Monday, September 19 at the start of class.

Week 3 September 21-26: The Normal Distribution, Sampling Distributions and Hypothesis Testing

- Read Howell Chapters 6 and 8
- Dougherty Review, Estimator properties
- Homework Assignment Three Due Monday, September 26 at the start of class.

Week 4 September 28-October 3: More on Hypothesis Testing

- Read Howell Chapters 12, 13.1-13.4, 14.1-14.4 9.1-9.9, 19.1-19.6
- Read Dougherty chapter 1
- Homework Assignment Four Due Monday, October 3 at the start of class.

Week 5 October 5-10: Exam 1 and Covariance/Correlation

- Exam one will be on Wednesday October 5
- Read Howell Chapters 8 and 19
- Read Dougherty chapter 1
- Homework Assignment Five Due Monday, October 10 at the start of class.

Part II Regression Analysis

Week 6 October 12-17: Simple Regression

- Read Dougherty chapters 2 and 3
- Homework Assignment Six Due Monday, October 17 at the start of class.

Week 7 October 19-24: Extending to Multiple Regression

- Read Dougherty chapters 3 and 4
- Redo midterm exam, no homework.

Week 8 October 26-31: Multiple Regression, a first look at a specification issue

- Read Dougherty chapter 4 and 7.1-7.3
- Read Hamilton chapter 7 on multicollinearity
- Homework Assignment Seven Due Monday, October 31 at the start of class.

Week 9 November 2-7: Dummy Variables & Interactions

- Read Dougherty chapter 6
- Read Thomas Brambor, William Clark, Matt Golder “Understanding Interaction Models: Improving Empirical Analyses” <http://polmeth.wustl.edu/retrieve.php?id=540>
- Homework Assignment Eight Due Monday, November 7 at the start of class.

Week 10 November 9-14: Transforming Variables

- Read Dougherty chapter 5.1, 5.2, 5.4, rest of chapter 7
- Read Hamilton chapter 7
- Homework Assignment Nine Due Monday, November 14 at the start of class.

Week 11 November 16 and 18: Exam and Heteroscedasticity

- Read Dougherty chapter 8
- Homework Assignment Ten Due Monday, November 21 at the start of class.

Part III Regression Extensions

Week 12 November 21 and 22: Autocorrelated Errors

- Tuesday is Friday, no class Wednesday or Friday for Thanksgiving holiday
- Read Dougherty chapter 13
- Turn in log file from Homework Assignment Eleven Monday, November 28 at the start of class.

Week 13 November 30-December 2: Models for Stationary Time Series

- Read Dougherty chapter 12
- Read De Boef and Keele
- Homework Assignment Twelve Due Monday, December 5 at the start of class.

Week 14 December 4-9: Limited Dependent Variable Models

- Read Dougherty chapters 11
- Read Hamilton Chapter 10
- Homework Assignment Thirteen Due Monday, December 12 at the start of class.

Exam 3 will be held during the scheduled final exam period.

Academic Dishonesty

The Department of Political Science, along with the College of the Liberal Arts and the University, takes violations of academic dishonesty seriously. Observing basic honesty in one's work, words, ideas, and actions is a principle to which all members of the community are required to subscribe.

All course work by students is to be done on an individual basis unless an instructor clearly states that an alternative is acceptable. Any reference materials used in the preparation of any assignment must be explicitly cited. IN an examination setting, unless the instructor gives explicitly prior instruction to the contrary, whether the examination is in-class or take-home, violations of academic integrity shall consist of any attempt to receive assistance from written or printed aids, or from any person or papers or electronic devices, or of any attempt to give assistance, whether the one so doing has completed his or her own work or not.

Other violations include, but are not limited to, any attempt to gain an unfair advantage in regard to an examination, such as tampering with a graded exam or claiming another's work to be one's own. Violations shall also consist of obtaining or attempting to obtain, previous to any examinations, copies of the examination papers or the questions to appear thereon, or to obtain any illegal knowledge of these questions. Lying to the instructor or purposely misleading any Penn State administrator shall also constitute a violation of academic integrity.

In cases of a violation of academic integrity it is the policy of the Department of Political Science to impose appropriate penalties that are consistent with University guidelines.

Disabilities

The Pennsylvania State University encourages qualified people with disabilities to participate in its programs and activities and is committed to the policy that all people shall have equal access to programs, facilities, and admissions without regard to personal characteristics not related to ability, performance, or qualifications as determined by University policy or by state or federal authorities. If you anticipate needing any type of accommodation in this course or have questions about physical access, please tell the instructor as soon as possible. Reasonable accommodations will be made for all students with disabilities, but it is the student's responsibility to inform the instructor early in the term. Do not wait until just before an exam to decide you want to inform the instructor of a learning disability; any accommodations for disabilities must be arranged well in advance.