1 Course Objectives

This course is designed to provide an introduction to quantitative political methodology for political science graduate students. We will cover basic probability and statistics through multivariate linear regression from a political science perspective. This means two things: that emphasis is placed upon political science applications and that what we will study reflects the dominant philosophies and practices of the discipline. Even if you yourself will not conduct quantitative (“large-n”) studies, you should be familiar enough with the basic techniques to be able to understand and critique them. This course accordingly aims to start you along the paths to becoming both a consumer and producer of quantitative research. It also serves as the gatekeeper to more advanced methods courses offered in the department, particularly to PS 206, which offers a more in-depth look at linear regression and its close relations. Since I not surprisingly subscribe to the well-known axiom that you only learn data analysis by doing, we will get our hands dirty from day one. Accordingly, you will become familiar with a major statistical software package: either STATA or R. Where possible, we will also read journal articles that illustrate or otherwise illuminate the conceptual and practical issues on the table.

2 Course and Contact Information

The syllabus, assignments, and other handouts are all available from the course GauchoSpace site. Announcements will be posted to the site; it should be your first port of call if you’re unsure about what’s happening when and where.

The best way to contact me is either to come to my office hours or to send me an e-mail.

- **Office Hours:** I encourage you to stop by early in the quarter so that you can get to know me and vice versa. Don’t, in other words, feel shy about coming to see me, whether you just want to chat about the weather or have more substantive course-related issues to discuss.

- **E-mail:** I will respond within twenty-four hours to e-mails that I receive during business hours on Mondays through Thursdays (I rarely check e-mail in the evenings). I will try to answer e-mails received prior to mid-afternoon on Friday by 5:00 p.m. but may not manage to respond until Monday.
3 Prerequisites

I assume only high school algebra and a tolerance for dirty hands. Regarding the former, getting your head around the material and making your statistical software package do what you want it to do can sometimes be both frustrating and time consuming. I can only encourage you to keep trying. With persistence (and sometimes a little help), you will eventually figure it out. Trust me on this—I’ve been there, too. Regarding the latter, calculus is helpful since some of the material we cover makes use of it. However, don’t panic: it is not necessary. I will show you some math from time to time in lecture, but this will be solely to provide you with motivating, behind-the-scenes intuition. You are not expected to be able to reproduce it in problem sets.

4 Requirements

Course requirements are five problem sets and a take-home final. Grades for the course will be calculated as follows:

- Five problem sets (10% each): 50%.
- Take-home final examination: 50%.

I strongly encourage you to form small study groups to work through the problem sets. However, the write-ups must be your own. This means that you can talk through how to solve a problem with someone else, but you must then on your own (in another room, later in the day, in silence) put the solution down on paper. No late assignments will be accepted without documentation. Extensions should be arranged at least twenty-four hours in advance.

5 Required Reading Materials

There is one required text for the course:


(henceforth, Gujarati). Equivalent chapters and page numbers from the Gujarati 2nd edition (1999) are shown in brackets. This is an older edition (a 4th edition was recently released), and the bookstore cannot obtain used copies. However, it can be ordered quickly from many used book vendors online, such as Amazon. That is accordingly the route that I recommend you take. If you are having trouble obtaining a copy, please come and talk to me about alternatives.

I will supplement this text with material from the following two even more user-friendly texts, a few copies of which will be made available in the graduate student computer room:


(henceforth, FPP) and

(henceforth, AF). Some of you may want to purchase one or both of these books, although it is not required that you do so. They are a source of elementary, clear, and almost completely math-free discussions of class topics. The Freedman text, in particular, will be very useful for those with little background in statistics. They also cover topics that may be of interest to you that the assigned texts do not cover.

Journal articles and other reading materials that appear on the syllabus can either be obtained in hard copy from the library or electronically from the web (for example, from JSTOR). To do the latter, you will either need to be on a university computer or to have configured your home computer to use the UCSB proxy server.

6 Recommended Reading Materials

Relative to the required texts, the following are of an approximately equivalent to slightly lower level of difficulty:


Also falling in this category are the FPP and AF texts referenced above. The text by Achen is a cheap, very accessible, and short overview of linear regression that is packed with a plethora of useful tidbits, one written by one of the greatest contemporary political methodologists. I highly recommend it.

For those desiring a more mathematically rigorous but still accessible approach to the material (i.e., some calculus and linear algebra), the following texts are good places to turn. The first, in particular, is widely regarded as one of the best textbooks on statistics ever written: from personal experience, I can say that it is well worth every penny.


Textbooks such as these are usually not cheap, but you will find yourself consulting the good ones throughout your academic career. Hence, view them as investments.

7 Computing

All of the assignments will require you to make use of a statistical software package to actually implement the methods that we discuss in class. You are allowed to use either STATA or R, the most popular packages amongst political methodologists.

STATA is available on computers in the department’s graduate student computing lab as well as in (by request) the Collaborate IT Labs, although the labs have a more recent version of the software. If you are interested in purchasing a discounted version of STATA for your personal use, let me know and I can set up a Course GradPlan. I recommend purchasing the Intercooled STATA 13 perpetual license, which costs $189. (Note that this is a more recent version than what is available in the computer labs on campus.)

R is the open source version of S-PLUS. (In the parlance of its developers, R is “not unlike” S-PLUS.) S-PLUS itself is unfortunately no longer available to us at UCSB since the university let its site license for the program lapse a few years ago. R, on the other hand, is completely free and can be easily downloaded from www.r-project.org. Instructions for downloading and installing it will be provided to you as part of the first assignment.

How do the two compare? Cutting edge methodological work in political science is increasingly done in R, and it is the computing package of choice of most statisticians. Further, R has graphing and data visualization capabilities that are only rivaled (and by some accounts outshone) by SPLUS’s. Both R and STATA offer command line interfaces instead of the graphical point-and-click interface that will be familiar to most of you from Windows. However, with STATA 11, STATA has added more of a point-and-click interface for the first time in its history. R’s commands are actually a programming language that is similar to C, which makes R both very intuitive for anyone who has ever studied programming and very powerful in that you can write your own ‘programs’ to run within R. (Its developers actually describe R as a ‘computing environment’ instead of statistical software to reflect this.) Hence, the preference of statisticians and political methodologists doing cutting edge work for R. R is arguably less well documented than STATA, though, and does not have pre-packaged implementations of some econometric models that are commonly employed in political science. Overall, while the start-up costs are relatively higher for R, those of you with a serious interest in political methodology are advised to give it a go. My primary expertise is in R, although I have learned the basics of STATA and use it from time to time in my own work.
8 Schedule

The tentative schedule for the course follows. Some topics will take more than a week, others less.

Introduction
Gujarati, Chapter 1 [Chapter 1].

Probability and Probability Distributions
Gujarati, Chapter 2 [Chapter 2.1–2.6].
FPP, Chapters 13 and 14.

Descriptive Statistics and Data Visualization
Gujarati, Chapter 3 [Chapter 2.7–2.8].
AF, Chapter 3.
FPP, Chapters 3, 4 and 7–9.

Important Probability Distributions, Sampling, and Limit Theorems
Gujarati, Chapter 4 [Chapter 3].
FPP, Chapters 16–19.

Statistical Inference: Estimation and Hypothesis Testing
Gujarati, Chapter 5 [Chapter 4].
FPP, Chapters 26 and 29.

Bivariate Linear Regression
Gujarati, Chapters 6, 7, and 9 (in part: 9.7–9.9) [Chapters 5, 6, 8.7; 9.8, 9.9 from 3rd ed.].

Multivariate Linear Regression
Gujarati, Chapters 8 and 10 [Chapters 7 and 9].