PS 3: Intro. to Empirical Analysis & QuantitativeMethods

Prof. Sean Gailmard
Dept. of Political Science
Office: 734 Barrows
Tel: 510-642-4677
Email: gailmard@berkeley.edu

Spring 2013
UC Berkeley
Location: VLSB 2050
Time: M-W 4:00 – 5:30
Office hours: M 2:00 – 4:00

This course provides an overview of some of the research methods that political scientists use to draw conclusions about the political world and public policy. What makes the subject matter "empirical" is that it deals with data and events that can be observed. The purposes of the course are to equip you to better understand research and policy debates you encounter in substantive political science classes, to sharpen your own ability to pose and answer research questions, and to impart skills in using some of these methods yourself.

When political scientists deal with data in research, the **two major issues are determining (*i*) what relationships and patterns occur in the data, and (*ii*) determining what causes the relationships**. Those are the two principal methodological problems we will confront in this course. We will use quantitative tools of statistics to ascertain relationships. We will explore several complementary methodologies — formal modeling,experimentation, quantitative (statistical) analysis — to assess causation in relationships.

Upon completion of the course you will be able to critique descriptive and causal statements made about empirical relationships in political science research, public policy analysis, and everyday life. You will also be able to apply basic tools of statistical description, modeling, and inference (such as regression and hypothesis testing) to a dataset.

A Note on Course Background: A large part of the course deals with formal and quantitative analysis, which is often expressed in mathematical symbols. However, the actual mathematical content of this course is fairly light and no math beyond high school algebra is required to succeed fully in this course. The problems students have more often come down to logic and grasping a style of thinking, than to mathematics per se. **It is my firm conviction that every student at Berkeley can fully succeed in this course.** If you find yourself struggling or need extra help, do not hesitate to let me know as soon as you recognize this. You will succeed if you are willing to think and work, but you have to help me help you.

Reading The source for most readings is the required text,*Fundamentals of Political Science Research* by Paul Kellstedt and Guy Whitten (New York: Cambridge University Press,2009). In addition, I may occasionally distribute required readings as PDF files.

GSI and Section Information The graduate student instructors (GSI's) are Stephen Goggin (head GSI, goggin@berkeley.edu), Jason Anastasopoulos (janastas@berkeley.edu), Michael Dougal (mdougal@berkeley.edu), Katherine Michel (katherine_michel@berkeley.edu),
Janna Rezaee (rezaee@berkeley.edu), and Sheryl Zaks (sherry.zaks@berkeley.edu).
All are Ph.D. students in the Political Science Department. The GSI's will teach sections, hold office hours, and mark the problem sets and exams.

Please note: You are not enrolled in the course unless youare enrolled in a section. If you are trying to add the course,you should not have any difficulty because there are still a few available seats. However, if the course fills to capacity and youwish to add it, attend the GSI section of your choice until your enrollment is resolved.

Grades The course grade will be based on an in-class midterm on Wednesday, March 6, a *comprehensive* final examination on Friday, May 17, problem sets, and participation. The course grade will be determined as follows:

- Midterm Exam: 25%
- Final Exam: 40%
- Problem sets (4 total): 30%
- Participation: 5%

I use a 90-80-70-60 scale for cutoffs of A, B, C, and D grades,with + or − designations in the top and bottom 2% of eachrange, respectively. A+ is a possible grade but is reserved for students whose performance is all-around exceptional; thus the determination of A+ is not strictly numerical and it is possibleto score 100% and not receive an A+. In practice about 1-2% of the class typically receives an A+.

I am not targeting any specific grade as the course average; if everyone can fully execute and master all techniques and concepts we cover, everyone will get an A. **Grades on exams and possibly for the overall course grade may be curved so that the average and range of the scores is reasonable.** What"reasonable" means exactly is context-specific and depends on how things go this term, but a curve can only raise scores, if applied (*i.e.*, I will never "curve down").

The exams will consist of a handful of short answer-type problems with multiple sub-parts each. Students are responsible for bringing blue/green books. Green books are preferred. Calculators will be allowed, but I will try to keep any arithmetic simple so it's easy to handle in your head. Calculators are allowed to do numerical calculations only, not to store parts of answers or definitions of concepts or any other exam aids in text storage facilities on your machine. GSI's proctoring exams will have full discretion to decide whether any calculators are being used in line with this policy, and to disallow their use if not.

The class participation grade depends partly on attendance in discussion sections: if you miss 3 or more discussion sections, you will earn 0 credit for participation. Beyond this, your GSI and I will jointly assess your participation in all facets of the course. We will construe it broadly, so don't feel compelled to ask questions in lecture every 10 minutes.

Attendance in lecture is not required, but two bad things will happen if you regularly skip: (1) your grade will suffer directly due to the participation score, (2) you will not understand the material as well as if you see the lecture material, particularly for the topics that the text does not treat in depth (which is, to be honest, a lot of the course).

Problem set due dates will be set throughout the term. Each of the problem sets will be weighted equally in the final grade. The problem sets will be posted on bSpace some time in the week before they are due. Problem sets are due at the beginning of lecture on their due dates. ***Electronic submissions will not be accepted*** because it is too easy for people to cheat and send me corrupted, un-openable files under the pretense of software incompatibilities. Late work will be penalized 20% per late day or portion thereof, and will not be graded if turned in more than 3 days late. Extensions will not be granted without proof of a legitimate medical or personal reason.

I encourage cooperation in study groups on problem sets. However,***the work you turn in for problem sets must be your own.*** Write up your own answers and do not copy from anyone else. **Plagiarism, and more generally, academic dishonesty of any kind, will not be tolerated.**

Sequence of Topics and Readings NOTE: K&W REFERS TO THE COURSE TEXT; OTHER READINGS WILL OCCASIONALLY BE DISTRIBUTED AS PDF FILES.

PART I: THEORY AND MODELS OF RELATIONSHIPS

Session1 Introduction. Course themes. The role of models in empirical research.
- K&W Chapter 1

Session2 Theory building in social science.
- K&W Chapter 2

Session3 Game theory I: strategic form; best response and Nash equilibrium; collective action problems.

Session4 Game theory II: collective action problems, examples.

Session5 Game theory III: electoral competition; the paradox of voting.

PART II: RELATIONSHIPS AND CAUSAL INFERENCE

Session6 Causation and causal inference in relationships. Potential outcomes model.
- K&W Chapter 3

Session7 Potential outcomes model, cont. Causal inference in experimental data.
- K&W Chapter 4 (through section 4.2)

Session8 Causal inference in observational data: Basic issues.
- K&W Chapter 4 (section 4.3 to end)

Session9 Causal inference in observational data: Advanced approaches.

Session10 Causal inference case study: Florida balloting in the 2000 election.
- Henry E. Brady, Michael Herron, Walter Mebane, Jasjeet Sekhon, Kenneth Shotts, and Jonathan Wand: "The Butterfly Did It: The Aberrant Vote for Pat Buchanan in Palm Beach County, Florida," *American Political Science Review* 95(4), 2001: pp.793-810. (Will be distributed electronically)

Session11 Review/Catch up

Session12 **Midterm Exam: Wednesday, March 6**

Part III: Descriptive Statistics — Understanding the Data You Have

Session13 Conceptualization and measurement.
- K&W chapter 5

Session14 Describing data: graphical and quantitative summaries.
- K&W chapter 6

Session15 Describing relationships: correlation and regression.

Part IV: Statistical Inference — Learning about a Population from a Sample

Session16 Conceptual foundations: Sampling, chance variability, distributions, and independence.
- K&W chapter 7

Session17 Sampling I: Expectation of sample means.

Session18 Sampling II: Standard Error of sample means.

Session19 Sampling III: Central limit theorem and normal distribution.

Session20 Hypothesis tests I: basics, Z test, and p values.
- K&W chapter 8

Session21 Hypothesis tests II: difference-in-means Z test.

Session22 Hypothesis tests III: examples.

Session23 Modeling relationships: bivariate regression.
- K&W chapter 9

Session24 Modeling relationships: multivariate regression concepts and examples
- K&W chapter 10

Session25 Multivariate regression II: examples
- K&W chapter 10

Session26 Catch-up and summary.

The final exam is Friday, May 17, 8:00 am - 11:00 am