

CONFLICT 811: RESEARCH METHODS I

FALL 2009

Professor Thomas E. Flores
Office: Truland 618
Phone: 703.993.9409
E-mail: tflores2@gmu.edu

Meeting Time: Thursday, 4:30–7:10 PM
Classroom: Truland 618
Office hours: By appointment and walk-in

Questions and Goals

“Establishing lasting peace is the work of education; all politics can do is keep us out of war.”
– Maria Montessori

“There are three kinds of lies: lies, damned lies, and statistics.”
– Benjamin Disraeli (often attributed to Mark Twain)

After a civil war ends in negotiated settlement, what makes it more or less likely that the civil war will recur? Does democracy promote peace by reducing military spending? Why do some countries grow economically during long bouts of civil war, whereas others collapse economically? Do treaties really modify their signatories' behavior and, if so, why? Are people who are patriotic more likely to be bigoted?

These questions are linked not only by their clear substantive importance, but also by our ability to answer them through statistical research. In fact, we'll treat each of these questions in this class. Statistical methods offer us the opportunity to rigorously compare cases across space and time; evaluate whether a hypothesis is generally true and/or only relevant in certain types of cases; engender new hypotheses for further consideration; and offer more rigorous advice to practitioners. More broadly, statistics creates excellent habits of mind in its practitioners, the most important of which is a relentless questioning of received wisdom, even when that wisdom is a cherished one. However, statistical research is accompanied by its own pitfalls, particularly when applied to social science: poor data that insufficiently measure the relevant construct; a lack of attention to the non-randomness of what we observe in the world (more on that later); analysis that is shoddy and opaque; the misapplication of a particular kind of model; and the deliberate distortion of otherwise sound analysis in order to mislead others or preserve one's cherished ideas. In short, the application of statistical methods to the scientific study of conflict presents both potential benefits and dangers.

Accordingly, this class is a doctoral-level introduction into quantitative reasoning and statistical research. It is intended as a companion to Conflict 812, *Qualitative Research Methods*. The goals of this class are two-fold, roughly corresponding to the two sections of the course. First, I hope to reinforce an in-depth understanding of some basic statistical principles, up to and including ordinary least squares (OLS) regression. By Week 8, you should know how to describe data statistically and graphically, understand basic sampling theory, estimate simple tests comparing two groups, and estimate a simple regression. Second, I hope to teach a few more advanced statistical techniques that are applicable to thinking about conflict analysis. Here, the goal is not to teach you the mathematical principles of those models, but their intuition and application, so that you can understand intuitively how an author, say, used logistic regression in an article on civil war.

More broadly, the goal of this class is not to produce statistical tool makers or even users, although if these tools appeal to you, I hope you will pursue your statistical training. Instead, this class is designed to train you to evaluate statistical work you read in journals and reports, interpreting that research carefully and understanding its strengths, weaknesses, and lessons for policy.

Fair warning: this class will likely be the most difficult class you take as a graduate student. We will have to move more quickly than any of us would like. Lecture and reading will require reading and thinking in mathematical symbols. However, the payoff is worth it. You will never think of the world in the same way after you take statistics.

Details: Requirements, Grading, Etc.

“I cannot give any scientist of any age better advice than this: the intensity of the conviction that a hypothesis is true has no bearing on whether it is true or not. . . . If an experiment does not hold out the possibility of causing one to revise one’s views, it is hard to see why it should be done at all.” - P.B. Medawar

Prerequisites

CONF 801, 810 and acceptance in the doctoral program. All others require my permission, which should be obtained immediately.

Course Materials

Readings for the course come in four forms. First, the following book is required. They are available for purchase at the Arlington branch of the GMU Bookstore (102 Arlington Blvd) and via online booksellers:

- Agresti, Alan and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences*. New York, NY: Prentice Hall. Fourth Edition. ISBN: 0130272957. ISBN-13: 9780130272959

Second, the following books are more generally recommended. Although I think the Agresti and Finlay is an excellent introduction, no one book is perfect for every student. The following books might be good for you if you seek to learn a particular set of skills.

- Abelson, Robert P. 1995. *Statistics as Principled Argument*. Lawrence Erlbaum Publishers. An excellent introduction into how to relate statistical results to your own research.
- Field, Andy. 2009. *Discovering Statistics Using SPSS* Sage Publications. I don’t use SPSS, but it’s still widely used. This looks like a good introduction to it. You could also check out *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows* by Julie Pallant.
- Hamilton, Lawrence. 2008. *Statistics with Stata 10*. Cengage Learning. Stata is one of the most-used statistical programs out there and this is a good introduction to how to implement many of the lessons we learn in this class using the program. For more on Stata, you can also try, *A Gentle Introduction to Stata*, by Alan C. Acock.

- Kennedy, Peter. 2008. *A Guide to Econometrics*. Wiley, John & Sons, Incorporated. A more rigorous approach to econometrics, the application of statistics to economics. Useful for those of you who are seeking a more rigorous textbook and/or are interested in economic issues in conflict analysis.
- Klass, Gary M. 2008. *Just Plain Data Analysis: Finding, Presenting, and Interpreting Social Science Data*. Rowman & Littlefield Publishers, Inc. An excellent introduction to how to interpret other people's data and present your own.
- Wonnacott, Thomas H., and Ronald J. Wonnacott. 1990. *Introductory Statistics*, 5th ed. John Wiley and Sons. A little dated, but still an excellent introduction to the study of statistics, on par with Agresti and Finlay.

Third, journal articles for the course are available electronically via our Blackboard site.

Fourth, I may assign a few print readings by passing out copies in class.

Participation

Given the rigor of the material at hand, the fast-paced nature of our course, and your status as Ph.D. students, I have high expectations of your effort level and performance in the classroom. Additionally, I think it's important to emphasize that you cannot prepare for a statistics class the way you might prepare for a discussion class, a lesson I learned quite painfully from personal experience. In an accompanying handout, I offer some recommendations as to how to prepare. However, let me offer some advice here. First, and most importantly, *do not fall behind — it will prove very difficult to catch up*. Second, complete all readings *before* the class in question. Class discussion and lectures will make no sense if you haven't read ahead of time. Third, I expect you to attend every class and arrive on time. Missing any section of class will set you back irreparably. Excused absences will be granted only rarely. Third, I will expect you to participate in the intellectual life of our class. I will often pepper you with questions as we work our way through the material and will choose respondents at random. If you are not prepared, I will know it. Participation can also come in the form of office hours attendance and/or electronic contact.

Assignments and Grading

There will be three components to your grade in this course.

- **Participation.** As discussed above, participating in class is required and will be worth **20%** of your grade.
- **Problem Sets.** You will be assigned four problem sets, with each one worth **10%** of your grade, or **40%** of your grade, cumulatively. The first will be due in class in Week 4 and cover Weeks 1 through 3, inclusive. The second will be due in class in Week 7 and cover Weeks 4 through 6, inclusive. The third will be due in class in Week 10 and cover Weeks 7 through 9, inclusive. The fourth will be due in Week 14 and cover Weeks 10 through 13, inclusive. More details will follow.

- **Final paper.** You will write a final paper (15-20 pages) that evaluates a researcher's statistical study of a topic in conflict analysis and resolution. In short, you will: choose a journal article or book article of interest to you and critique the data and statistical methods employed by the author. The paper will be worth **40%** of your grade. More details will follow.

I know that many of you are concerned about your grades and I will do everything in my power to help you throughout the course. However, I do have high standards for you, a function of the respect I have for your ability and ambition. Due to that respect, I simply will not allow you to produce work that is below your potential. Therefore, do not expect a high grade for minimal effort.

Software

The first two problem sets will at times require you to use computer software. This is line with the goals of the first part of the course, one of which to build your ability to work with data. Given the computing resources available at ICAR, I cannot require you to use a statistical package to complete those problems, though I encourage you to purchase such software if you are contemplating further study in statistics. However, I will assume that you have access to MS Excel or a similar spreadsheet program. In many cases, the spreadsheet portion of Google Docs — which is free — will suffice. If you are interested in using statistical methods in the future, then you should look into acquiring either SPSS or Stata, or plan to spend time in our computer lab while you work on problem sets and the final paper. I believe your class experience will be augmented by access to statistical software.

Course Outline

“Man is capable of producing more complex behavior than he is capable of understanding. The behavior of an infant baffles a psychologist, and vice versa.”

- Charles Lave and James G. March

Course Structure

As discussed earlier, this class has two goals, which correspond roughly speaking to the two halves of the class outlined below. In the first half of the course, we will introduce (for some of you) and/or review (for others) several core statistical concepts. I refer to these techniques as ‘core’ because without a firm understanding of them, it would be impossible to learn more about either regression (Weeks 6 and 7) or any of the more advanced topics in Part II. In those classes, we therefore will blend statistical theory (*i.e.*, math on the blackboard) and applications thereof. The summit of our Part I climb is Weeks 7–9, when we take up regression analysis, the workhorse model of statistics.

Part II changes in tone and focus. Here, we study six statistical issues that, for various reasons, simple regression does not tackle very well. Each of the issues is particularly important for the study of social science. Rather than tackling the technical aspects of the model, we will study the models more practically. We’ll begin with the intuition of the issue at hand and then shift to a real-world example drawn from recent statistical research.

Part I. Reinforcement: From Descriptive Statistics to OLS

Week 1 (9/3). Introduction: A Philosophy of Science. Or, Why Use Statistics?

Week 2 (9/10). Measurement, Descriptive Statistics and Graphics

Week 3 (9/17). Probability and Distributions

Week 4 (9/24). Statistical Inference: Point Estimation

Week 5 (10/1). Statistical Inference: Significance Tests

Week 6 (10/8). Two-Group Comparisons and Measures of Association

Week 7 (10/15). Bivariate Regression

Week 8 (10/22). Multiple Regression I

Week 9 (10/29). Multiple Regression II

Part II. What To Do When OLS Won't Do

Week 10 (11/5). Dichotomous and Nominal Dependent Variables

Week 11 (11/12). Time Series and Hierarchical Models

Week 12 (11/19). Endogeneity 1: Non-Random Selection

Week 13 (11/26). No Class: Thanksgiving Break

Week 14 (12/3). Endogeneity II: Joint Endogeneity

Week 15 (12/10). Factor Analysis

Class Structure

Given the nature of learning statistics, class will be structured in a specific way:

- Each class will begin with logistics — handing in assignments, taking roll, etc.
- We then will take about an hour to discuss the theory of the topic at hand, mostly from the textbook. During that time, I first will emphasize the intuition of the subject material. Then we will move on to the mathematics of the material, translating the intuition into the math and back again. If we have time, we'll take a five-minute break.
- Next, we'll take an hour to demonstrate the theory with reference to a real-world statistical example. For the first eight weeks of class, all examples will be drawn from one article, an analysis of military spending by Benjamin Goldsmith. That will provide some continuity to the class, as we'll spend several weeks concentrating on one problem conceptually: why do some countries spend more than others on their militaries? During the second week of class, we'll focus on a different scholarly article each week. If we have time, we'll take a five-minute break.
- Finally, we'll spend 20 minutes highlighting the next week's material. This will give you an opportunity to get some insights into the next week's class and hopefully facilitate your reading.

Detailed Course Schedule

Part I. Reinforcement: From Descriptive Statistics to OLS

Week 1. Introduction: A Philosophy of Science. Or, Why Use Statistics?

Required

- Lave, Charles and James G. March. 1975. *An Introduction to Models in the Social Sciences*. Pages 1-7; 10; 21-29; 40-42; 52-60. Skim pp. 61-78 if desired.
- Singer, J. David. 1968. "The Incomplete Theorist: Insight Without Evidence." Chapter 4 in Singer. 1968. "The Incomplete Theorist: Insight Without Evidence." Chapter 4 in Klaus Knorr and James Rosenau, eds. *Contending Approaches to International Politics*. Princeton, NJ: Princeton University Press. pp. 62-86.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton, NJ: Princeton University Press. Chapter 1 (pp. 3-33).
- A&F, Chapter 1.
- Goldsmith, Benjamin E. 2003. "Bearing the Defense Burden, 1886-1989: Why Spend More?" *Journal of Conflict Resolution* 47(5): Introduction (551-552) and "Competing Hypotheses" (554-560).

Week 2. Measurement, Descriptive Statistics and Graphics

Readings:

- A&F, Chapter 2 (Section 2.1 only, pp. 11-14)
- A&F, Chapter 3 (pp. 31-71)

Week 3. Probability and Distributions

Readings:

- A&F, rest of Chapter 2
- A&F, Chapter 4 (pp. 73-105)

Week 4. An Introduction to Statistical Inference: Point Estimation

Readings:

- A&F, Chapter 5 (pp. 107-141)

Week 5. An Introduction to Statistical Inference: Significance Tests**Readings:**

- A&F, Chapter 6 (pp. 143–181)

Week 6. Two-Group Comparisons and Measures of Association**Readings:**

- A&F, Chapter 7–8 (pp. 183–253)

Week 7. Bivariate Regression**Readings:**

- A&F, Chapter 9 (pp. 255–299)

Week 8. Multiple Regression I**Readings:**

- A&F, Chapter 10–11 (pp. 301–367)

Week 9. Multiple Regression II**Readings:**

- A&F, Excerpts from Chapters 12–15 (TBA)

Part II. What To Do When OLS Won't Do**Week 10. Dichotomous and Nominal Dependent Variables****Readings:**

- A&F, Chapter 15 (pp. 575–619)
- Hartzell, Caroline A. 1999. "Explaining the Stability of Negotiated Settlements to Intrastate Wars." *Journal of Conflict Resolution* 43(1): 3-22.
- DeRouen, Jr., Karl R. and David Sobek. 2004. "The Dynamics of Civil War Duration and Outcome" 41(3): 303–320.

Week 11. Time Series and Hierarchical Models**Readings:**

- A&F, Sections 16.1–16.2
- Re-read Goldsmith, Benjamin. 2003. “Bearing the Defense Burden, 1886–1989.” *Journal of Conflict Resolution* 47(5): 551–573.
- Curtis, James E., Douglas E. Baer, and Edward G. Grabb. 2001. “Nations of Joiners: Explaining Voluntary Association Membership in Democratic Societies.” *American Sociological Review* 66(6): 783–805.

Week 12. Endogeneity I: Non-Random Selection**Readings:**

- Excerpts from Achen, Christopher. 1987. *The Statistical Analysis of Quasi-Experiments*. Berkeley, CA: University of California Press.
- Von Stein, Jana. 2005. “Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance” *American Political Science Review* 99(4): 611–622.

Week 13. No Class: Thanksgiving Break**Week 14. Endogeneity II: Joint Endogeneity****Readings:**

- Excerpts from Achen, Christopher. 1987. *The Statistical Analysis of Quasi-Experiments*. Berkeley, CA: University of California Press.
- Kang, Seonjou and James Meernik. 2005. “Civil War Destruction and the Prospects for Economic Growth.” *The Journal of Politics* 67(1): 88–109.
- A&F, Section 16.4

Week 15. Factor Analysis**Readings:**

- A&F, Section 16.5
- de Figueiredo, Jr., Rui J. P. and Zachary Elkins. 2003. “Are Patriots Bigots? An Inquiry into the Vices of In-group Pride.” *American Journal of Political Science* 47(1): 171–188