THE POLITICAL METHODOLOGIST

NEWSLETTER OF THE POLITICAL METHODOLOGY SECTION
AMERICAN POLITICAL SCIENCE ASSOCIATION
VOLUME 10, NUMBER 2, SPRING, 2002

Editor: SUZANNA DE BOEF, PENNSYLVANIA STATE UNIVERSITY
sdeboef@psu.edu

Contents

Notes From the Editor

This issue of *TPM* features articles on teaching the first course in the graduate methods sequence and on testing theory with empirical methods. The first course presents unique challenges: what to expect of students and where to begin. The contributions suggest that the first graduate methods course varies greatly with respect to goals, content, and rigor across programs. Whatever the nature of the course and whether you teach or are taking your first methods course, Chris Achen's advice for students beginning the methods sequence will be excellent reading.

With the support of the National Science Foundation, the *empirical implications of theoretical models (EITM)* are the subject of increased attention. Given the empirical nature of the endeavor, EITM should inspire us. In this issue, John Londregan and Rebecca Morton share their thoughts on "methods & modeling". Be sure to note the announcement of the first EITM training institute to be held this summer at Harvard University. Deadlines for applications are very soon.

This issue also contains 3 book reviews and a discussion of issues related to selecting software packages for the analysis of time series data. In the "LaTeX Corner", Andrew Martin introduces LaTeX for Unix users. Section news includes many alternative ways to spend your summer expanding your toolkit as well as announcements from our president. Jonathan Nagler invites proposals to host the 2003 and 2004 summer methods meetings and announces the beginning of the search for the next editor of *Political Analysis*.

Thanks to all who contributed to this issue. Please contact me with your suggestions and ideas for future issues of *TPM*.

Suzanna De Boef

Teaching (and Learning in) the First Graduate Methods Class

Teaching Introductory Probability Theory

Michael Herron
Northwestern University
m-herron@northwestern.edu

My experience with political science graduate students suggests that a first class in statistical methods should teach essentially no statistics. Instead, it should focus almost exclusively on the basics of probability theory, a subject about which most beginning doctoral students in political science know almost nothing. Probability theory is the foundation for inference, and without some basic knowledge of it many important concepts in statistics are hard to grasp. One can, of course, learn how to apply statistical procedures and perhaps even interpret output without knowledge of probability theory. I have found, however, that students who do not know any probability invariably understand very little about statistics, even if they are proficient at implementing sophisticated procedures.

Why is probability theory so important? Consider hypothesis testing, a topic common to many quantitative (and conceivably qualitative) research designs. Hypothesis testing involves the question, "Given my theory, did I observe data that was expected or data that seems very unusual?" Without drawing on probability theory, clarifying "very unusual" in a careful way is next to impossible.

Similarly, sampling distributions, which are fundamental to statistical inference, cannot be understood without some rudimentary knowledge of what a random variable is. But, grasping the definition of a random variable requires knowing about distributions, and distribution theory, even at its most basic level, is based on simple events, sample spaces, and so forth, i.e., on the building blocks of probability theory.

Ideally, of course, first year graduate students in political science would begin their doctoral studies having already taken a year of calculus, basic probability, linear algebra, and so forth. Barring this rosy scenario, my approach to introductory statistical methods teaching at Northwestern University has been to spend approximately an academic quarter (ten weeks of classes) on basic probability theory as preparation for a second quarter class on regression analysis. At Northwestern, all political science doctoral students, except those majoring in traditional political theory, are required to take the introductory class that I teach as well as the following regression class.

The text I have traditionally used in my one quarter course on probability is *Probability: An Introduction*, by Samuel Goldberg. This book was published in 1960 and remains in print as a Dover edition. Thank goodness for Dover! Goldberg's text is a true gem, and it is a book that is accessible by statistics beginners whose background in mathematics is not overly strong. Moreover, Goldberg can also challenge students who are comfortable with the concept of proof and who do not need to review basic concepts like set theory.

Goldberg is a wonderful book, one well–suited to a basic probability course, for a number of reasons. For starters, Goldberg does not require any calculus. Even so, it is very rigorous and Goldberg waves no hands, so to speak. All claims in the text are proved and, regardless of how hard they try, students who work from Goldberg will not be able to dodge proofs. It is very important to me that students who finish my introductory course in probability theory can distinguish between what can be called a result (something that can be proved starting from axioms) and a finding (something discovered with data). Feel free to choose your own vocabulary here; the difference, in this context, between "result" and "finding," and why students should understand this difference, is obvious. Goldberg, since it is entirely theoretical, offers students time to become accustomed to the concept of a result.

Goldberg develops probability theory for finite sample spaces, and starting with basic set theory everything is derived from first principles. For example, Chapter 1 reviews basic set theory and notation; I have learned from experience that this is very important material and it is useful to spend a class on it. Students who have trouble with unions and intersections will generally not be able to understand collections of events, sample spaces, random variables, distributions, and so on.

Chapter 2 introduces the notion of probability and the idea that a probability is a function defined on a sample space. I have found that students occasionally have trouble defining a probability as a function. Nonetheless, Chapter 2 covers a number of key concepts: probability of a union, probability of a complement, conditional probability, Bayes's Rule, and independence, just to name a few. All of these topics are important, but I believe that conditional probability and Bayes's Rule are two of the most important. Understanding the difference between $P(A|B)$ and $P(B|A)$ is key, and the difference is generally easy to motivate with students. Focusing on a court example, where $P(guilt|evidence)$ needs must be distinguished from $P(evidence|guilt)$ is often helpful.

In addition, one can generally find an example of getting conditional probabilities backward in publications like *Newsweek*, if not in the occasional job talk. Many journalistic articles on, say, prisons, will invariably conclude that having incarcerated parents increases the probability of a child's spending time in prison. This conclusion, notably, does not follow from a visit to a prison and the observation that most prisoners have parents who, at one time or another, had spent time prison.

After covering basic probability I spend a week on counting rules, permutations, and combinations. I have found that teaching combinatorics is always a stretch because many students inevitably question why they have to know how many ways there are to pick eight senators such that none is from California and only one is from Illinois. But, counting problems can be made quite fun and this helps overcome their lack of immediate relevance. Of course, counting is essential to understanding binomial random variables and, if one does not understand binomials, most distributions are opaque.

Goldberg has many excellent combinatoric problems, but the vast majority have little to do with political science. So, when studying counting rules I usually make up my own problems that have obvious political connotations.

The most significant challenge for me when teaching introductory probability theory is bridging the gap between distributions on finite sample spaces, e.g, the Bernoulli distribution, and continuous distributions, e.g., the normal distribution. Eventually, after developing random variables, moments, probability functions, and distribution functions, my class leaves Goldberg because I have to teach something about normal random variables. My approach when encountering continuous distributions like the normal distribution has been to teach by analogy.

For example, the mean of a random variable with finite support is a weighted average, and this is something that Goldberg covers explicitly. Indeed, when my class works through expectation, problem sets involve writing down probability functions, directly calculating means, and so forth. This emphasizes the importance of expectation as an average. Then, because integrals are limits of sums, I can draw on what my students know about expectation to give intuition about the mean of a continuous random variable.

Similarly, when distribution functions are introduced, they are accompanied with full details and are defined on finite sample spaces. This allows students to calculate directly these functions, to plot them, to understand how they differ from probability functions, and so forth. Then, when the normal distribution function appears, I am able to draw attention to the parallels between it and a distribution function for a discrete random variable.

One of the main difficulties that I have encountered with my probability class is students who are impatient with theory and who want to dive immediately into applications. My class has no computer component, and it is obvious up front that no data will be analyzed in it. I have taken two approaches to impatience with theory. The first is to emphasize continuously throughout the quarter the material that is coming next in an applied regression class. This may not convince everyone; noting that random variables are important because regression estimates are themselves random variables may not be useful to students who have never seen regressions.

The second thing I do is to make sure that almost all of my examples and problems are social science based. Goldberg, in this context, is not that helpful as many of the problems in his book are either abstract or have science/engineering applications. Nonetheless, my running examples throughout the quarter involve democracy and war (how many democratic wars are enough to discount the theory of a democratic peace?) and voting in presidential elections (how are polls interpreted?).

Because Northwestern attracts many qualitatively oriented graduate students, I make a special effort to emphasize how the principles of probability theory apply regardless of whether one's dissertation is quantitative or qualitative. There are obvious limits here, as a student who is purely qualitative and does not want to develop any fluency in quantitative methods probably will not want to learn about random variables. But, topics like conditional probability and Bayes's Rule apply to all empirical research designs, and I believe that students should hear this over and over.

If I had extra time for my probability class, i.e., if Northwestern were on semesters rather than quarters, I would probably spend additional classes on continuous distributions and teach about them using simulations. This would presumably build intuition—students could see the law of large numbers and the central limit theorem in action—and would also teach some programming skills. I have tended to avoid introducing computing in my quarter course on probability so that I do not have to spend a large amount of time on computer skills.

What my colleagues and I want to do at Northwestern is attempt to ensure that students, during their required methods training, learn some fundamentals of probability theory and, just as importantly, come to believe that "F7 Econometrics" is a bad way to conduct research ("F7 Econometrics" refers to a propensity to rely on computers for statistics. . .). Statistical software is becoming easier and easier to use, and this is certainly a boon to all. But, one downside of readily accessible software is that it can foster an attitude in which fundamentals are not important. Spending a quarter on probability theory will help prevent this and will also give students a foundation for future learning.

First Things First (whatever "first" happens to mean): Syllabus Choices for Statistics I

Eric Plutzer
The Pennsylvania State University
exp12@psu.edu

Teaching introductory statistics shares many challenges with other quantitative classes. One unique challenge, however, is dictated by its position as first in a *sequence* of classes that are intended to build cumulatively. Syllabus choices (the type of textbook, the sequence of topics, and the relative emphasis on various skills and experiences) have implications for how students will fare in, and how my colleagues will teach, intermediate quantitative methods.

Before sharing my opinions and experiences, however, it is useful to consider the nature of cumulation and I turn to the inventor of the cumulative scale.

> Consider a mathematics test composed of the following problems:
>
> (a) If r is the radius of a circle, then what is its area?
>
> (b) What are the values of x satisfying the equation $ax^2 + bx + c = 0$?
>
> (c) What is de^x/dx?
>
> If this test were given to members of the [APSA], we would perhaps find it to form a scale for that population.... The responses to each of these questions might be reported as a dichotomy, right or wrong. There are $2 \times 2 \times 2 = 8$ possible types [but] *for this population* we would probably only find four of the possible types occurring. There would be the type that would get all three questions right, the type that would get the first and second questions right, the type that would get only the first question right, and the type that would get none of the questions right. (Guttman 1944, 143, emphasis added).

Guttman's example illustrates the properties of what would later become known as a "Guttman scale." But for pedagogy, a more useful insight occurs later in his article (p. 149):

> An interesting problem associated with scales is: why do... [items] form a scale for a given population?... There is no necessary reason why a person must know the area of a circle before he can know what a derivative is, and in particular the derivative of e^x. The reason for the scale emerging is largely cultural.

That is, geometry is *usually* taught in tenth grade, algebra in eleventh, and calculus in senior AP courses or in college. Sequences that seem "natural" are not necessarily the only way, or the best way, of helping students learn. So the way that *I* learned statistics is not the best or only way to do so - nor is the way that you are I currently teach statistics. But this is not a lesson I learned right away.

The first time I taught the introductory course I looked for textbooks utilizing the same sequence of topics, level of mathematical rigor, and philosophy as my own intro course 15 years earlier. Those of my and earlier generations will recognize the "Blalock" sequence (also found in texts by Hays, and the Wonnacotts), which was characterized by:

1. No calculus or matrix algebra

2. A thorough introduction to probability theory in the first third of the course

3. A "classical" approach to hypothesis testing (reject or fail to reject the null, disdain for p values and "nearly significant")
4. Emphasis on explained variation as the core of descriptive statistics
5. Frequent hand calculations supplemented by a few (4-6) computer assignments.

Teaching it the "natural" way, I discovered that the "Blalock" approach had three important implications for a department's *statistics sequence*. First, the Statistics II instructor becomes responsible for teaching "math for social scientists" (calculus, matrix algebra, and complex distributions). This often makes the second class seem "harder" or the second instructor seem to be more of a "quantoid." Yet even without advanced math, the extended treatment of probability early in the course alienated students with math anxiety and/or limited prior exposure to quantitative methods (few burning political issues are illuminated by the distinction between combinations and permutations or by hand-calculating binomial distributions – jury selection being an important exception that I use frequently). Third, the traditional texts give minimal attention to linking statistics to issues of causal inference and research design, thereby requiring supplemental texts.

Dissatisfied with the approach that had previously seemed "natural", I sought alternative approaches for the following year, and these seemed to fall into five categories.

The "applied statistics" approach is best exemplified by Bohrnstedt and Knoke's top selling text. Designed to focus on interpretation and intuition while exposing students to many techniques, a fifteen-week course using their text can cover elementary descriptive statistics, hypothesis testing, multiple regression, plus log-linear, logit and LISREL models! To me (despite the book's popularity and my friendship and intellectual debts to Bohrnstedt), this seems best suited for a class of policy students who will only get one statistics class and need to read policy literature and interact with applied statisticians.

Most econometrics texts (e.g., Ramanathan) assumed that students had already completed an undergraduate course in probability and statistics - an unrealistic assumption for many graduate programs in political science. This left three approaches. One is simply to skip "introductory" material altogether and start with a low-level regression text. At Iowa State, graduate students successfully used Bowerman and O'Connell's *Linear Statistical Models*, which "reviews" means, standard deviations, probability, hypothesis testing and differences of means in two short chapters (Gujarati's *Essentials* employs a similar approach, though without matrix algebra). This is tempting because you get into multivariate models quickly and the approach and notation blend well with intermediate econometrics texts used in the next class. But I decided against this approach because students who lack an outstanding undergraduate statistics course are left to understand foundational material on their own.

This leaves two familiar approaches and both presumably would cover similar material - but in a very different sequence. One approach assumes that later coursework will build on *mathematical* foundations; the other assumes that later coursework will build on students' "feel" for quantitative data, description, and an understanding of how research design, data, and the "real world" generating the data cohere. Mindful of Guttman's insight, and in the absence of empirical research on the effectiveness of each approach, the choice is largely a matter of taste, instructor strengths and weaknesses, and coordination with colleagues teaching the next class.

The first option devotes most of the first semester to "math for social scientists": the algebra of expectations, set theory, probability, distributions of commonly encountered random variables, and at least enough calculus to understand how OLS normal equations are derived. Armed with a strong mathematical foundation, students can move relatively quickly through many statistical applications and the colleague teaching the second class in the sequence can select a high-end econometrics textbook. As students encounter more data sets in their graduate careers, the mathematical foundations get fleshed out.

To be honest, I did not think I could do a good job with this approach. Perhaps projecting my own biases, I concluded that this would turn off students with little exposure to quantitative methods who nevertheless had great potential as empirical researchers.

Instead, I adopted what I now call the *Bauhaus* approach, as it shares the architectural school's credo that *immersion in raw materials* is the essential foundation for creative problem solving.

I wanted (1) to provide students with substantial immersion in data and descriptive data analysis, (2) assign more than a few hand calculations of means, deviations, squared deviations, variances and other raw materials that are manipulated to generate t-tests, regression slopes, covariance matrices and their derivatives (e.g., SEM, measurement models, etc.), while (3) helping students see the correspondence between theories and expected empirical outcomes.

Most textbooks in the "Stat 101" market adopt the first goal, by immersing students in descriptive statistics: univariate descriptives, comparisons of means and proportions, and regression in the first portion of the class — but *without* any discussion of statistical inference or stochastic processes (see Janda, previous issue of *TPM*). *After* students are comfortable describing many different data sets, these texts cover probability and hypothesis testing. Most such texts, however, are far too simplistic and move too slowly.

I discovered that good graduate students could work through a rigorous undergraduate statistics text at two chapters per week (I've used Devore & Peck, and Anderson & Finn). This meant that we complete all of the traditional "Blalock" topics (albeit in different order and with added emphasis on description), and learn a statistical package, in nine weeks. This then allows $2\frac{1}{2}$ weeks to discuss the logic of causal inference - emphasizing that causal theories require certain patterns of data if they are valid (I rely on tabular elaboration but simulation could also be employed to good effect here) - and $2\frac{1}{2}$ weeks to work through a basic multiple regression primer (I use Lewis-Beck's Sage monograph, supplemented by Allison's *Primer*).

I'd like to think that most students complete the class being more interested in *quantitative data* than they were before, and that the ability to construct, read, and interpret complex regression tables gives students a feeling of accomplishment that motivates future research, and an eagerness to work through the math for social scientists that is now the burden of one of my colleagues. Indeed, I'd *like* to think that this is the best and only rational way to introduce students to quantitative methods.

But thanks to Guttman's insight, and the excellent results of colleagues taking a completely different approach, I now know better.

References

Allison, Paul D. 1998. *Multiple Regression: A Primer.* Pine Forge Press.

Anderson, Theodore W. and Jeremy D. Finn. 1996. *The New Statistical Analysis of Data.* Springer.

Blalock, Hubert M. 1979. *Social Statistics.* New York: McGraw Hill.

Bohrnstedt, George W. and David Knoke. 1994. *Statistics for Social Data Analysis.* F.E. Peacock.

Bowerman, Bruce L. and Richard T. O'Connells. 1990. *Linear Statistical Models.* Duxbury.

Devore, Jay L. and Roxy Peck. 1993. *Statistics: The Exploration and Analysis of Data.* Duxbury.

Gujarati, Damodar. 1992. *Essentials of Econometrics.* McGraw Hill.

Guttman, Louis. 1944. "A Basis for Scaling Qualitative Data." *American Sociological Review* 9 (Apr): 139-150.

Hays, William L. 1994. *Statistics.* Harcourt Brace.

Lewis-Beck, Michael S. 1980. *Applied Regression: An Introduction.* Sage Publications: Thousand Oaks, California. Dryden.

Ramanathan, Ramu. 1998. *Introductory Econometrics With Applications.*

Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. *Introductory Statistics.* Wiley.

Teaching Graduate Statistics: Reflections from a Small and Diverse Program

Lawrence J. Grossback
West Virginia University
Larry.Grossback@mail.wvu.edu

Teaching the first graduate statistics course can be a challenge. Teaching it in a small department that combines masters and doctoral students in both academic political science and professional policy analysis brings that challenge to new heights, heights I was not initially prepared to scale. I came to the graduate methods sequence recalling the soothing advice of my adviser. After I expressed some concern about teaching up to four methods courses, she calmly reminded me that I had just recently completed these courses myself, and that as long as I took good notes, the courses were well on there way to being prepped. No such luck. I took very good notes from an excellent instructor, and they have done me almost no good. My notes are well suited for a class of doctoral students intent on becoming empirical political scientists. I find myself teaching graduate students from a diverse set of majors, who have diverse professional goals, and who mostly come with little mathematical or statistical training. I imagine that like me, many new methodologists find themselves in such a situation, and so with a cautionary note that I am relatively new to the game, I offer some reflections on my early efforts.

Integrating Research Design and Empirical Analysis

We offer a typical four course methods sequence: research design, statistics, regression, and advanced empirical topics. This often leads to two things. First, students dread the empirical classes to come. Second, our students learn regression a year into the program, well after taking substantive seminars that require them to read empirical research. I have dealt with this by addressing empirical research early in the research design seminar. I begin with a discussion of the emergence of empirical methods and the reasons why empirical evidence is valued in political science and policy studies. I present this apart from a discussion of the philosophy of science, not to discount opposing views, but to ensure that early in their careers, students understand why we demand they spend a significant amount of time learning this material.

Several other things have helped me teach students to gain more comfort with and interest in empirical methods. The first book I assign in the design course is Berry and Sander's primer on regression, *Understanding Multivariate Research*. They designed this book to guide students in reading regression results before having taken the class, and in my experience it works exceedingly well. After students read this book, I have them write a short paper that requires them to find an article of their choice that uses OLS and to then recreate the regression table and explain each part. I follow this with a methodological critique of an example of empirical analysis, currently Carmines and Stimson's *Issue Evolution*. Students report that this early focus on regression provides a good deal of comfort when approaching empirical research in other seminars. Finally, I have found that the focus on descriptive and causal inference in King, Keohane, and Verba's, *Designing Social Inquiry*, is a helpful setup for the statistics class, but perhaps even more valuable in the statistic and regression classes. In the later courses, I have students reread sections of the book, and they report this helps them better understand both research design and the logic of inferential statistics.

Focusing on Statistical Reasoning and the Big Picture

The idea that technical skill and basic understanding are not the same things might be a cliché, but at least it is an accurate one. An early focus on inference and the use of sample information can help students with the underlying logic of inferential statistics, but for those with little statistical training, it leaves basic but important concepts such as expected value, variation, probability, and significance levels to the more technical treatments of statistics texts. A colleague and I have both had trouble building a broader understanding of the logic behind these concepts and we have found that this, more than technical concerns, can impede students when they move on to regression. My colleague suggested I use Cuzzort and Vrettos's book, *The Elementary Forms of Statistical Reasoning*, as a supplement to our text. The focus of the book is on linking basic statistical ideas to everyday forms of reasoning and learning. The book does an admirable job of linking basic topics to key ideas that play a role in later, more demanding, sections of the course. For example, it nicely links central tendency and averages to developing expectations and later to issues of probability and sampling. I found that this helps build a better understanding of the logic behind empirical analyses, and that students pay more attention to a subject when they see why it is relevant to future topics.

I have made the big picture a key element of my statistics course. To me the big picture is applied statistics and the successful completion of the regression course. I begin the course with an introductory chapter from a more advanced regression text in order to show students where we are heading in terms of estimation theory and hypothesis testing. I then return to the big picture to introduce topics, especially those that require some instruction in mathematics and which can be seen by students as doing math for the sole reason of making their lives difficult. For example, before a series of lectures on set theory, probability, and constructing probability distributions, I offer a brief overview of the role of probabilities in drawing samples, evaluating estimators, and conducting hypothesis tests. My students still report suffering through the math, but they find some comfort in knowing why they are suffering.

An Early Focus on Measurement Issues

Part of understanding the big picture of empirical analysis includes coming to an understanding of the importance of measurement. Political methodology courses have not always done this topic justice, with many offering only a cursory assessment of validity and reliability. I add my voice to the call for a greater focus on measurement mainly because I have found it to be a valuable tool in teaching new students the benefits of empirical analysis, the pros and cons of both qualitative and quantitative methods, and how to critically evaluate research. I currently use sections from Nunnally and Bernstein's (1994) psychometric theory text that address the value of quantification and measurement. The focus on the benefits of standardized measures fits nicely with other efforts at explaining the value of empirical methods. More recently, political scientists have produced helpful articles. Adcock and Collier (2001) offer an accessible look at the history of validity assessment and the processes of validating measures. Jacoby (1999) offers an updated look at the importance of understanding the levels of measurement and how measurement decisions are theoretical statements about the political world that are subject to falsification. I combine these readings with an exercise that requires students to critique the measures used in an article by assessing multiple forms of validity and offering a plan to test the reliability of the measurement techniques used. Dealing with measurement issues early in the first statistics course has helped me make the point that the quality of all empirical analyses rests on the quality of the measures used. A lesson sometimes lost in the focus on conducting hypothesis tests and running regressions.

Engaging Students and Offering Feedback

In conclusion, I offer two final suggestions. First, I have found it helpful to engage students in either analytical exercises or research problems early and often, sometimes at the expense of excluding problem sets. Because my students come with little research or empirical training, I find it easier to keep them interested in technical topics if they are working with statistical packages and real research questions. Many of my assignments allow students to work with articles or data in their fields of interest. This appears to better hold their interest and allows them to develop a sense of how empirical methods are used in areas where they intend to do research. Finally, I try to finds ways to get students feedback from other faculty members, especially those in areas of political science where my own knowledge is lacking. One tool we have had success with is a poster session where students in the research design course present their final research designs. The participants praised the experience and were pleased with the feedback and suggestions for further research. The poster session had the added benefit of introducing new graduate students to the entire faculty and serving as a means to involve students in other conferences both on campus and off. A strong interest in research goes a long way toward ensuring their success in the first statistics class, mainly because their early sense of dread can be replaced by a new interest in having the tools to answer research questions.

References

Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(September): 529-546.

Berry, William D. and Mitchell S. Sanders. 2000. *Understanding Multivariate Research: A Primer for Beginning Social Scientists*. Boulder: Westview Press.

Carmines, Edward G. and James A. Stimson. 1989. *Issue Evolution*. Princeton: Princeton University Press.

Cuzzort, R.P. and James S. Vrettos. 1996. *The Elementary Forms of Statistical Reason*. New York: St. Martin's Press.

Jacoby, William G. 1999. "Levels of Measurement and Political Research: An Optimistic View." *American Journal of Political Science* 43(January): 271-301.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.

Nunnally, Jum C. and Ira H. Bernstein. 1994. *Psychometric Theory*, 3rd Ed. New York: McGraw-Hill.

A Stealth Approach to Quantitative Methods: Getting Students to Use Quantitative Methods in Their Own Research

Charles Tien
Hunter College and Graduate Center, CUNY
ctien@hunter.cuny.edu

When I was asked to teach the graduate methods course at the CUNY Graduate Center I gladly accepted. It would give me an opportunity to teach and get to know graduate students. Many students expressed fear about methods and some even appeared to loathe the required course. Many of these feelings were projected onto me, the instructor. The welcome I received was chilly at best from a few of the students who appeared to have the most anxiety about the course.

Some resented having to take a required course on a topic that they considered irrelevant to their interests in political science. They saw the course as an onerous obstacle to getting their degrees, something to overcome and of little use to their own research interests. It was the math and statistics many had run away from as undergraduates. Most political science graduate students do not go to graduate school to learn statistics or quantitative research methods, they go because they love politics. And many are surprised at the amount of attention and coursework devoted to research methods in their programs. This seems especially true at the CUNY Graduate

Center where many of the students are primarily interested in areas that use less quantitative methods than the American politics subfield. Remembering back to my own first semester graduate school experience, when I went to purchase my books I thought the bookstore had mis-shelved them as I did not find many political science sounding titles, but instead founds books like Freund's Mathematical Statistics and a calculus textbook. Goals for teaching the first graduate quantitative methods course often must account for these student fears.

I have two goals when teaching the first quantitative methods course to graduate students. My first goal is to get them to put aside their fear and strong dislike of the topic by mid-way through the semester. My second goal is to get them to think about using quantitative methods in their own research by the end of the semester. I do not try to turn them *into* quantitative research methodologists but rather turn them *onto* quantitative research methods. Of course, I am delighted when one or two students do take an interest in methods and continue to pursue it. But my hope for the majority of the students at the CUNY graduate program is that they will eventually use quantitative methods in their own research. If I can get students to use it in their own work, then their interests for more advanced methods will be self-imposed and these students, I think, are more likely to pursue additional methods courses.

I use a stealth approach to try to get students to overcome their fears and eventually think about using quantitative methods in their own work. The stealth approach includes weekly homework assignments using datasets chosen by the students themselves. In other words, my hope is that without really knowing it some students will become excited about methods by the end of the semester, and will want to pursue it because they think their research will benefit from it. I believe that a lot of the anxiety that students have about the first quantitative methods course comes from a lack of confidence in their ability to understand the material. Therefore, I try to build up their self-confidence with weekly homework assignments. Weekly assignments alleviate fear by getting students to approach the subject in small steps, and it gives them (hopefully) positive feedback. Once students realize that they can actually do the work there is less reason to fear it. If I assigned less homework or only a mid-term exam, I think the students' anxiety levels would increase as students are more likely to put off doing the hard and gradual work that is necessary to grasp the material. If students only have to focus on interpreting slopes, intercepts and significance tests in one week, the task seems more manageable as opposed to having to study half a semester's worth of material for a mid-term exam. After two months of doing weekly assignments, students should develop enough confidence to successfully complete the course.

To get students to think more about using quantitative methods in their own research, I require them to find datasets that are related to research topics of interest to them. Then I require students to use their own data in the weekly homework assignments. I encourage students to scan the ICPSR holdings for a dataset that is related to their own substantive areas of interest. This requires the cooperation of a patient ICPSR campus representative, as numerous student requests for data all at once can be quite burdensome. Now that ICPSR is allowing direct data downloads from computers physically located on member campuses, I anticipate that acquiring data in a timely manner will be much easier for students. As students do homework assignments on univariate, bivariate and ultimately multivariate statistics over the course of the semester on data that is interesting to them, I find that many become interested in wanting to know more, not only about their data, but also about methods. For example, multiple regression assignments often result in students wanting to analyze models with dichotomous dependent variables, which requires them to go further in methods and take a course involving logit and probit analyses. Having students working on different datasets also saves me from having to read assignments on one topic over and over again. For example, students have used data from National Health Interview Surveys, the Euro-Barometers, and various education surveys in addition to the American National Elections Studies data.

My joy in teaching methods comes from providing some of the same tough love I received from my own methods professors. I consider research methods as good medicine for graduate students. I tell them it will make their research better. By the end of their first quantitative methods course, I hope that students will take what they have learned and apply it to areas of research that they are interested in. If they start to think of research in an empirical framework, then I believe that they are one big step closer to becoming quantitative researchers themselves.

Advice for Students Taking a First Political Science Graduate Course in Statistical Methods

Christopher H. Achen
University of Michigan
achen@umich.edu

For many graduate students, the study of elementary statistics is a demanding experience. Often, the course challenges their stamina and professionalism like no other course they have taken previously. The style of thought is unfamiliar to them, and its no-nonsense structure may appear arid and inhumane. Their happy undergraduate life, full of effortless adolescent success and idealistic speculation, is now seen to be lost forever. The kindly and indulgent undergraduate faculty who convinced them to study political science have now been replaced by unsympathetic and unforgiving researchers, who insist that they adopt the new alien language and its cold thought. If the course is taken in the fall, as often happens, then for these students, the term proceeds painfully slowly through ever colder days, and as December approaches, a wintry season overtakes both mind and body. The instructor and the unhappy students stagger on, neither enjoying each other, until the holidays finally release them both.

Or so goes the mythology. Actually, no such outcome is inevitable. No matter what one's level of preparation, there is no reason to undergo a painful experience in basic statistics. Reasonable care and effort can produce an experience that is, if not pleasant, at least comfortably endurable and professionally profitable. Of course, as with crossing streets or cooking over an open fire, bad experiences do occur, and the victim is not always to blame. But many accidents are due to carelessness, inattention, or self-indulgence. Plain good sense will give most students a satisfying experience in basic statistics.

If you are beginning elementary statistics yourself, the first point to realize is that you should tailor your course planning to your individual needs. Graduate students take statistics courses for all sorts of reasons. If your area of interest is Machiavelli, for example, you may want to take a statistics course just to read the political science journals and to follow what some of your faculty and fellow graduate students are doing. You may have no desire or necessity to master everything in the course; an informal level of understanding may suffice. No sensible person will ever care what your grade is. Relax and enjoy it!

In fact, acquiring this intuitive understanding of statistics should be a goal for every student in the course, not just for humanistic students, but even (or especially) for those students who plan to specialize in quantitative political science. At the intuitive level, the student learns the vocabulary, the style of work, and sorts of questions addressed by statistical methods. Not every "statistical finding" in the newspapers or the professional journals is reliable, and not every important topic in political science can be addressed with statistical techniques. A course in quantitative methods helps sort out the quantitative and the humanistic, and within the quantitative realm, it should aid in distinguishing the true from the false, and the researchable from the unknowable. In short, at the intuitive level, the student becomes a knowledgeable reader and consumer, with sound substantive judgment about what is worth doing with quantitative techniques. No student of political science, whatever the field, should be without at least a little skill of that kind.

Most graduate students in the elementary course, however, will have professional needs that require them to go beyond sound substantive judgment about data and intuitive understanding of inference, important as those are. They will also require a knowledge of applied elementary statistics. That means acquiring a working grasp of the basic theory and a little actual experience of doing statistical research. It also means getting past mechanical use of canned computer packages and developing an understanding of when their output should be believed. For students in this group, whether they intend to do quantitative work themselves, or merely read, judge, use, and teach the results of those who do, some personal experience of doing the work themselves is needed for their professional futures.

If this is your situation, you should recognize that your background is probably quite unlike that of the student next to you. No course in political science graduate training programs treats a greater range of student preparation than does the elementary statistics class. Some students will have good mathematics backgrounds and perhaps even prior work in undergraduate statistics; others will have forgotten all their high school algebra. The same learning strategy will not work for both. Thus you need to tailor your course planning to your preparation.

Even among students who have taken no quantitative courses since high school, circumstances differ. Some students may just need a brush-up. Others will be at a more severe disadvantage. I have had students who could not remember whether, if A = B and B = C, then does A = C? For them, the course will be difficult, perhaps sufficiently so that they should take a refresher course in high school math first. If you find yourself in that position, do not confuse your lack of coursework with lack of ability. If knowing the statistics is important to you, do not try to skip steps and get by on grit (or belligerence). Instead, go back and do what your fellow students have done: take the prerequisites. When you return to statistics, you will be amazed at how much easier the material will have become and how much faster you will learn it. The time you "lost" will be gained back.

Most students, though, have adequate preparation and are ready for the course. Elementary statistics courses in political science departments are aimed at the average mathematical background, and most students will find themselves with adequate groundwork. If you belong to this larger group of students, you can focus on learning the new material in the course. For that, however, you will need help. Lectures and homeworks are designed to provide it. Homework problems are critical, and much of your learning will occur as you do them. Working in groups can also be helpful, but don't use your group as a crutch and let other people do your thinking for you. Better a few C's on the homeworks and an A on your first professional research paper than vice-versa.

The textbook is meant to help you learn, too. Ah, the textbook. You will almost certainly dislike the text—virtually every student does, no matter which book is chosen. Most of the problem is that quantitative thinking is not a large part of most undergraduate political science courses, and so students come to elementary statistics with learning skills that translate poorly to a scientific context. For example, students may read only 300 pages of basic statistics in an entire semester, while they may be assigned up to several thousand pages in their other graduate courses—equivalent amounts of reading. Overlooking that, they allocate half an hour for reading 20 pages of statistics, as would be more than adequate in their other courses. When understanding at that rate proves impossible, students decide that the book is poorly written, and perhaps also that the course is "too theoretical."

No one can read mathematical material in the same way that one reads history or novels. Patient, line-by-line study is needed, pencil in hand. Sometimes an hour goes by on a single page. Sometimes one has to make up problems for oneself before a point is truly understood. Too often, students do not know this. They have gotten by with memorizing in previous mathematics courses and never learned to truly understand. If you find that this is your situation, the advice is the same as my professor gave me thirty years ago: find a quiet place to study, with a hard chair and a good light. Allocate enough time for the reading, and learn to read in the new way. This is easier than it sounds. Most of the challenge is seeing that one needs new learning skills; once you seek them, the skills themselves will arrive relatively quickly.

All that said, sometimes the text will stump you. No text works well for everyone, and no text works well all the time for anyone. Be aggressive about finding a companion text that suits you and that gets you past difficult passages in the main text. Ask your professor for tips about other texts at the same level as the one you are using in class. There are dozens of introductory statistics texts in your college library. Spend a couple hours going through them during the first week or two of the course, and find one that works for you. Those two hours will save you half a week's work later on. Keep the book handy the rest of the term, reading it as needed.

Yet another supplemental book can be helpful as well. In any mathematical field new to me, I like to read a seriously dumbed-down book first, just to get the feel of the subject. Such books are often well written verbally, but they may have mathematically sloppy arrangements and slightly wrong intuitions that will make your professor cringe. Never mind! You'll forget all the mistaken details eventually anyway. Get the big picture in mind so that you have a feel for what you are doing and where the course is going, then fill in the details from the regular textbook so that your research work is right.

For this purpose, ask your professor to recommend "good, short, chatty books written at much too low a level for this course," perhaps books that would be used for undergraduates. There are dozens of such introductory statistics texts at a variety of levels, all the way down to picture books. Find one that works for you in the course you are taking. Don't wait until the end of the course to read it, when your confusions will have built upon each other and work pressures will have accumulated. Get it read the first two or three weeks of the course.

Above all, don't expect immediate success if you have been away from mathematics for awhile. You have work to do. Don't start ignoring the texts, letting your colleagues do your homework problems, and expect to be spoonfed by the lectures. The lectures will help, but in a course like basic statistics, slothfulness is fatal. You need to improve your intuitions by working partly on your own, doing both reading and problem sets. That way, you can hear the lectures and read the text with a firm foundation of previous material and an intuitive understanding of where the presentation is going and why. In turn, that will make the mathematics much easier. In this course more than most, steady work is rewarded.

Lastly, a word to those students for whom the class will expose previously unsuspected talents and interests in quantitative work. For you, the class will turn out to be intellectually fulfilling, perhaps even fun. It will open the way to additional coursework in political methodology and formal theory, and that in turn will lead to a lifetime of professional success and intellectual satisfaction. In a stealthy way that you may not notice immediately, the course will change who you are.

As it dawns on you that you are in this group, you will see that you need to understand the subject more seriously than we can teach you in the introductory course. Why, for example, do we estimate the mean and median of a normal distribution with the sample mean, while we estimate the mean and median of a double exponential distribution with the sample median? When is maximum likelihood estimation a good idea, and when does it produce a foolish estimator? To answer questions like these, you will need to learn enough calculus and linear algebra to take several further courses in political methodology and econometrics, as well as additional coursework in mathematical statistics. Some aspects of statistical theory are important to formal theorists as well, and will be taught in game theory courses. The introductory statistics course may open up all these worlds to you.

But even if you are in this group, you, too, have much to learn from the usual political science introductory course. Political methodologists and formal theorists are not professional statisticians, and it is important not to get lost in the mathematics and computing to the exclusion of political data and political understanding. Don't ask the introductory course to replace a rigorous introduction to mathematical statistics. That you must learn elsewhere. But do ask the introductory statistics course to connect you to the right political topics, topics where the mathematics and the data can be intelligently deployed. Then go learn the math you need, and come back to political topics to do some science.

In summary, with a willingness to learn, a little hard work, and a certain maturity of spirit, the introductory statistics course can be a rewarding experience for nearly all students. That is not to say that it will be easy. (Indeed, if most members of the class are finding it easy, their future careers are probably being sacrificed to temporary comfort.) The point is rather that, for those students working in quantitative areas of the discipline, successful completion of this course takes them to a milestone on a road to professional competence. That is why, if you are beginning such a course, careful planning is so important. You need to assess both where you are starting from and which professional road you are on. With those decisions made, the trip through introductory statistics, challenging though it may be, can bring great professional satisfaction.

Testing Theory

Political Theory and Political Reality

John Londregan
Princeton University
jbl@princeton.edu

Perhaps the great divide between methodologists and the rest of the profession (and indeed, most of the rest of the human race) is our ingrained tendency to build models that incorporate a stochastic element. This has profoundly important implications for the way that we relate our models to data.

Incorporating randomness into our models instills a healthy aversion to "anecdata" and to "argument by counterexample". While we have all heard such arguments from our colleagues, let me illustrate the missleading use to which anecdata[1] can be put in an apolitical context. My grandmother, may she rest in peace, lived to be 84 years old, and smoked at least a pack of cigarettes daily. A typical "argument by counterexample" would use this as evidence that smoking did not shorten one's life span. Of course, readers will impatiently note that we don't know how long my grandmother would have lived had she not smoked. Defenders of "argument by counterexample" might then contend that we can never know what would have happened, and that tests must be based on observable outcomes. Fair enough a methodologist might respond, but we should look for a systematic relationship based on extensive datasets. How does the survival of cigarette smokers compare with what we would expect to see from an otherwise comparable set of non-smokers?

While the desire for a large sample is intuitive, it stems from our reliance on the error term. When the stochastic component is negligible, a single case can make or break a theory. The perihelion of Mercury, first photographed (Dicke 1967) in 1919, was more consistent with

[1] An idea for future issues of TPM—a small and suitable reward for the person who comes up with the best name for the units in which anecdata should be described—"story", "case", and other words fail to capture the malleability and divisibility of units of anecdata such as the outbreak of the first world war.

Einstein's general theory of relativity than with its competitors, and persuaded many physicists to take general relativity theory seriously (Skinner 1969). While subsequent confirmatory evidence helped to solidify the case (Taylor 1979), physicists were largely persuaded to take the theory seriously on the basis of a single datapoint.

Ultimately, we require more observations because our models leave considerable leeway for randomness. An unintended but significant side effect of our incorporation of randomness into our models is a greater tolerance for misspecification. A typical argument goes as follows: the linear or log-linear-quadratic specification we are using is a Taylor's series expansion of an unknown functional form, with the approximation error (we hope innocuously) rolled into the error term. Of course, many non-methodologists will balk at accepting a linear specification (or any other concrete functional form). Essentially it is an article of our faith that baptism by inclusion of an error term washes clean the original sin of misspecification.

Of course, we don't have much choice–fully flexible functional forms lead to the "parameter proliferation problem", while any concrete choice of functional form is almost certainly a misspecification. But specifying a functional form that we are virtually certain is wrong raises the question of what we are really doing when we test a hypothesis. In general, tests of substantive models of political behavior are simultaneously tests of ancillary assumptions about functional form, and about the distribution of the random error terms in our model. Thus, if we should reject the null, we do not know whether rejection resulted from the model being wrong, or from flaws in what we already know is a misspecified functional form that we have used to operationalize our model. Similar problems arise throughout the social sciences, for a careful discussion in the context of finance see Roll (1977). Not only are we left uncertain how to interpret rejection of our models, there is the additional question of what we should make of acceptance when we are virtually certain that we are using a misspecified functional form.

The two examples at the beginning of this essay, the first from biology, the second from physics, suggest that there are different kinds of models. There are the "squishy" models of biology that identify "risk factors" and "propensities" and the closely specified models of physics, with precise predictions and razor's edge rejection criteria. Of course, if we were to apply the criteria of the physicists and build our models from the microfoundations up political scientists would first need to solve the outstanding problems of sub-atomic chemistry and molecular biology, the political actors we study being made up of atoms and cells. The messier models of biology lack the nicely trimmed edges of their counterparts in physics, but they can be very practical nevertheless–a case can be made that more lives have been saved by the discovery that cigarettes promote lung cancer and heart disease than by the theory of relativity (and the nuclear weapons whose design it makes possible). This in spite of the raggedness of biological models, and their various, often implicit, *ad hoc* assumptions about functional forms and probability distributions.

On a continuum of models, from the tightly specified theories of physics, to the partial models of biology, with their tattered and untied loose ends, we operate much closer to the biological end of the spectrum. This means that our models are all misspecified, and so, "wrong" as literal descriptions of the world. But they can never-the-less be very useful precisely because they simplify a complex world in a way that is understandable, and yet retains an important element of the underlying causal relationships among the variables we are interested in.

When we calibrate our models we are learning about orders of magnitude and about the "closeness" of our models to the data, which is to say, to the world we are trying to understand. Since our models are misspecified, it is absurd to think that hypotheses tests are actually telling us about the "truth" of our models–we already know that they are false. Rejecting a hypothesis in a dataset is really a statement about the closeness of a misspecified model to our data. Pursuing useful misspecified models of a messy reality is not as satisfying as a quest for the holy grail of "true theory". But let's not forget that it was a messy misspecified model with a low R^2 that identified cigarette smoking as a risk factor for cardio-pulmonary disease, a finding that has dissuaded a substantial number of people my age and younger from smoking, and so, in expectation, saved the lives of some of the people who read this article.

References

Dicke, Robert Henry. 1967. "Gravitational Theory and Observation." *Physics Today* 20: 55–70.

Roll, Richard. 1977. "A Critique of the Asset Pricing Theory's Tests." *Journal of Financial Economics*, 129–76.

Skinner, Ray. 1969. *Relativity.* Blaisdell: Waltham.

Taylor, J.H., L.A. Fowler and J.M. Weisberg. 1979. "Measurement of General Relativistic Effects in the Binary Pulsar 1913+16." *Nature* 277: 437.

EITM*: *Experimental* Implications of Theoretical Models

Rebecca Morton
New York University
rebecca.morton@nyu.edu.

While most political scientists are aware of the growing influence of formal modeling within the discipline and the increasing emphasis on the empirical evaluation of this work, the experimental approach to empirical study of formal models is still rare and not well understood. In fact, I suspect if you asked an "average" political scientist about experimental research within the discipline he or she would answer with a focus on the work of political psychologists such as Milton Lodge, Diana Mutz, Shanto Ivengar, Paul Sniderman and the more recent field experiments of Don Green and Alan Gerber. While I certainly do not mean to argue that all of this work is atheoretical, there can be no argument that this work is generally not testing what we think of as solved formal models of the sort addressed by the initiative coming from the NSF.[1] A few might know about the work evaluating formal models with experiments, but are less likely to see this work as particularly influential or to be able to name some noteworthy experiment within this research tradition as they would from the political psychology tradition. In graduate courses on research methods or substantive topics within political science, to the extent that students read and discuss experimental papers, they are generally political psychological experiments as well. Experimental work evaluating formal models generally is relegated to a part of a specific course on the experimental methods or on formal modeling, although the latter is rare. [2]

This raises two related questions:

1. Why is this the case?
2. Are experiments useful in evaluating theoretical models within political science?

I will take these questions in reverse order, first explaining why I think experiments are useful for empirical evaluation of formal models and then why it is that few seem to recognize the value of that work.

Why and how use experiments to evaluate formal models?[3] When we engage in devising formal models two things happen: we make assumptions that are either known to be false or whose verification is unknown and we derive predictions from these assumptions. The assumptions we make, since most of the time, as social scientists, we are studying individual behavior, are assumptions about individual behavior whether as single decision makers or in groups. We assume voters, candidates, party elites, defense and foreign ministers, etc., maximize expected utility or some variant, we make assumptions about how individuals use and process information, we make assumptions about the extent that individuals consider the choices of others in their decision making, calculate probabilities, etc.

The assumptions are one of the principal complaints that non formal theorists make about the approach broadly and often are much of the discussion among formal theorists when comparing models. They are extremely difficult to evaluate using naturally occurring data since the empirical model that would be used to evaluate the assumption would contain a host of ancillary assumptions from which it is virtually impossible to disentangle from the evaluation of the assumption. In contrast, when working in the laboratory, a researcher can use experimental controls to isolate the details of the particular assumption of concern. In Morton (1999), chapter 5, I provide some examples of how experiments have been used in this context to empirically evaluate the assumptions underlying expected utility theory and the separability of preferences, both prime aspects of many formal models in political science. Frankly, to the extent that we can evaluate the behavioral assumptions we make in formal models, I know of no other way to do this than experiments and it is incumbent on any researcher who works on the empirical implications of formal models to value and appreciate the use of experiments for assumption evaluation.

Nevertheless, most empirical evaluation of theoretical models in political science, since it is mainly nonexperimental, focuses on predictions of the models. Since the true value of theory is to derive predictions from the underlying assumptions, an emphasis on prediction testing is expected. However, experiments are extremely useful for this type of work as well. I believe that empirical evaluation of theoretical models requires (at some point) that the empirical model used is as close as possible to the theoretical model to be evaluated.[4] Otherwise, the

[1] Green and Gerber (forthcoming), however, have explicitly argued that their approach is atheoretical and that this is the advantage of their work.

[2] The equating of experiments in political science with political psychology is evident in the fact that my own institution's recently established interdisciplinary Center for Experimental Social Science represents political science solely by the research area of political psychology and the type of researcher sought to fill the first joint hire between the Center and my department was a political psychologist.

[3] Much of my views on these issues is already published [Morton (1999)]. In that book, I try to provide numerous examples of experimental evaluations of formal models, across the subfields of political science, which I do not have the space to review here. Here I will try to offer a summary of some of the main points.

[4] Some would argue that such evaluation should always be "seamless" between theory and empirical research, however, I do not hold

evaluation is not of the formal model but some mixture of the formal model and the ancillary assumptions made in devising the empirical model. The empirical model when it is not " "seamless" with the theory, as my coauthor Charles Cameron and I note (forthcoming), is really just implied by the formal model, not derived from it. If the empirical model's predictions do not hold up with compared with the data, we don't know if this is a rejection of the theory as an acceptance of the empirical model cannot mean that theory is accepted as well.

There are ways to devise empirical models, which are equivalent with the original formal model, but generally this means making the model more complicated, which can hide the empirical implications and sometimes be unsolvable. Charles Cameron and I, in our review of the literature, found only a few such works because of these two realities. However, in the laboratory we can work with the model in its simpler form, using the control of experiments to fit the empirical world more closely to the theoretical world rather than the other way around. We generally contend that when the model is not supported in the controlled environment, then it is highly unlikely it is supportive when confronted with naturally occurring data (or if such support occurs it is because of the disconnect between the formal model and the empirical one). This type of experimental analysis of formal models is called "theory tests" and most experimental research on formal models within political science takes this approach.[5]

But what if the formal model is supported in the laboratory, surely we cannot jump to the conclusion that it explains the naturally occurring world, so what use is this work? The value then comes from what can follow using the experimental method. Experiments can allow the researcher to individually relax the assumptions of the formal model, conducting what we call "stress tests" of the theory, finding out the limits of the theory in a scientific and careful manner. For example, an experiment testing the median voter theorem may relax the information subjects have about candidate positions that is assumed in the theory, while holding constant the information that candidates have about voters. This type of control is available when we make statistical assumptions in a empirical model using naturally occurring data, but it is limited by truthfulness of those assumptions. The laboratory provides an extremely clean method of moving from the theoretical world to the naturally occurring world, which is not possible when you move in the other direction except when we apply statistical assumptions.

that view as will become evident in the rest of this essay. Theory evaluation that is only seamless has limits just as empirical evaluation which is never seamless does as well.

[5]See Davis and Holt (1993) and Morton (1999) for discussions of theory and stress tests.

Finally, the laboratory allows us to evaluate theory, which is simply not possible to evaluate in the naturally occurring world because of a lack of data. When my coauthors and I (Gerber, Morton, Rietz (1998), wanted to test theories of how cumulative voting systems worked, we were able to generate a large number of elections using that system and compare the system with standard plurality rule holding preferences of voters constant, and since cumulative voting systems have been used very rarely and little data of this sort is available when they have been used, the empirical evaluation could not have taken place without the laboratory.

So, given the value of experiments in evaluating formal models, why is it that few political scientists are attracted to the method or think of it when we discuss experiments in the discipline? I suspect the main reason is the ambivalence that most political scientists feel about artificiality, as exemplified in the criticism of formal models within the discipline and the NSF initiative to fund work on the empirical implications of formal models. Because laboratory experimental research on formal models typically works so closely with the theory, researchers unfamiliar with the work have difficulty seeing it as "real" tests of the theory since the environment is controlled. What gives these types of experiments their advantage is seen as the principal reason for ignoring them.

Political psychologists have similarly found that work that is highly controlled is less influential than research that loosens the control (increases the supposedly external validity of the empirics). It is no surprise that political scientists almost immediately embraced in courses and citations the results from the acknowledged atheoretical Gerber and Green field experiments on voter turnout, even though few can tell us what these experiments add beyond the research of Rosenstone and Hansen on the same issue. In fact, the rise in influence of experiments from political psychology, I suspect, is mainly due to the new methods (such as laboratories at malls, computer generated realities) that lessen the concerns political scientists have about external validity. Moreover, the empirical research on formal models using naturally occurring data (working with maximum likelihood functions) resembles (although the differences between the two are profound) the use of close to theory free empirical methods to explore political science data, so the constraints of and role played by the theory are less evident to the reader.

The trouble with the bias of political scientists against the type of experiments, which can be used effectively to evaluate formal models, is that the choice does not need to be so stark. There is a continuum between an experiment with high internal validity (which the theorist values) and one with high external validity (which the empirical researcher wants, like field experiments) and if we encouraged experimentation we can begin to move between these extremes in meaningful and productive ways.

Political scientists are finally beginning to recognize that formal model does not mean rational choice (and what we mean by rational choice is highly variable as well), which is much to the discipline's good. But to dismiss or relegate laboratory experiments on formal models to a little discussed specialty because they are seen as mainly theory tests and thus not useful, ignores the reality that these controlled evaluations allow us to move along this continuum in productive and multidimensional ways not possible if we only use naturally occurring data. Moreover, the bias towards emphasizing external validity eliminates two principal advantages of experimental evaluations of formal models that are, I argue, virtually impossible using naturally occurring data - assumption evaluation and empirical research on counterfactuals (which we can only speculate on outside of the lab). If we are truly going to begin an agenda that focuses on the empirical implications of formal models, experimental research should take the primary role and political science would become a "real" science.

References

Cameron, Charles and Rebecca Morton, (forthcoming), "Formal Theory Meets Data", in *State of the Discipline of Political Science, Volume III,* edited by Ira Katznelson and Helen Milner, NY: W.W. Norton & Co.

Davis, Douglas D. and Charles A. Holt.1993. *Experimental Economics*, Princeton: Princeton University Press.

Gerber, Alan and Donald Green (forthcoming), "Reclaiming the Experimental Tradition in Political Science", in *State of the Discipline of Political Science, Volume III*, edited by Ira Katznelson and Helen Milner, NY: W.W. Norton & Co.

Gerber, Elisabeth R., Rebecca B. Morton, and Thomas Rietz. 1998. "Minority Representation in Multimember Districts", *American Political Science Review* 92 (March): 127-144.

Morton, Rebecca, 1999, *Methods and Models: A Guide to the Empirical Evaluation of Formal Models in Political Science*, Cambridge: Cambridge University Press.

Rosenstone, Steven J. and John Mark Hansen, 1993. *Mobilization, Participation, and Democracy in America*, New York: Macmillan Publishing Company.

Articles

The LaTeX Corner: LaTeX For the Rest of Us

Andrew D. Martin
Washington University
admartin@artsci.wustl.edu

Have you ever tried to send a Word document to a colleague that uses a different version of Word? Or even worse, to one that uses Linux, or, God forbid, a Macintosh? If so, you will remember how things seem to change across every version of Word across every platform. Your page fifteen might be your colleague's page twelve and so forth. In a world with growing use of Linux (and other Unix-based operation systems), and with a resurgence of the Macintosh platform (which – perhaps not coincidentally – is now too a Unix-based operating system), the ability to share and collaborate on documents across platforms has become increasingly important. There are file formats, such as postscript and Adobe PDF that accomplish this goal for finished-product reasonably well. But for works-in-progress, these solutions are quite limited. One of the distinct advantages of using LaTeX as a text-processing system is its seamless ability to move from platform to platform with no changes. This is based in part to the platform-independent TeX implementation by Donald Knuth, and a commitment among the LaTeX community to maintain cross-platform compatibility.

In the last issue of *The Political Methodologist* Chan H. Nam wrote a very nice introduction to LaTeX for the uninitiated. In that article he outlines how to get started with LaTeX, and details some implementations of LaTeX for the Windows operating system. Because of its cross-platform friendliness, his introduction applies equally well for those using Macintosh or Linux. In this article, I will point to some valuable resources for using LaTeX on my operating systems of choice – MacOS X (and it's younger sibling MacOS) and Linux. That is, this article serves as a summary for the enlightened who choose to avoid Microsoft Windows for their computing needs.

LaTeX on Linux and other Unices

As Linux (and other Unices) become a more and more cost-effective desktop system, more and more political scientists will begin using Linux not only for statistical computation, but also for day-to-day computing. Installing and configuring LaTeX on a Linux system by hand is a rather cumbersome process. But nearly all of the major Linux distributions include the `tetex` package (`http://www.tug.org/teTeX/`). This package provides all you need to turn your .tex source file into a DVI file (usually you just need to type `latex myfile.tex`). Another nice feature in this package is the ability to directly generate PDF files instead of DVI files; one does this by typing (you guessed it) `pdflatex myfile.tex`. The resulting PDF file can be read by the standard utilities.

Not only are LaTeX .tex source files platform independent, so too are DVI files. Indeed, you can preview a DVI created on any system on any other system, *and you will see precisely the same thing*. The Linux utility to preview DVI files is called `xdvi`, which is also contained in nearly all of the standard Linux releases. You can download `xdvi` from `http://www.math.berkeley.edu/$\sim$vojta/xdvi.html`. One can convert a DVI file to a postscript file using the handy `dvi2ps` utility, which is part of the `tetex` distribution.

I cannot conclude my discussion of LaTeX on Linux systems without mentioning text editors. Of course, you can use *any* text editor to type and edit your .tex source files, but there are some tools out there that make the job much easier. Many people swear by Emacs (`http://www.gnu.org/software/emacs/`, the "extensible, customizable, self-documenting real-time display editor." Emacs is distributed under the GNU GPL, and is available for nearly every operating system (including Windows). There is a package called AucTeX(`http://mirrors.sunsite.dk/auctex/www/auctex/`) that makes writing LaTeX code as easy as possible in Emacs. Others prefer a more visual approach. I can recommend the editor AlphaTK (`http://www.santafe.edu/$\sim$vince/Alphatk.html`) that runs in the X-Windows environment. It has a superb graphical interface that makes entering LaTeX easy, especially for "hard" things like tables, lists, and mathematics.

LaTeX on the Macintosh

From my perspective, one of the most exciting things in the world of computation is MacOS X. This new operating system has one thing in common with the old MacOS – an intuitive, efficient, and aesthetically pleasing graphical user interface. The nuts and bolts of MacOS X, however, are completely different. In fact, MacOS X is built upon an Open Source version of BSD Unix called Darwin. If you like the command line as much as I do, one can now install the standard Unix tools for LaTeX on a Macintosh running MacOS X. You can run a X-Windows server on top of the Macintosh GUI, and use the same `xdvi` as you would on any other Unix machine.

But MacOS X offers much more to the user than the standard Unix command line tools. And LaTeX on MacOS X is no exception. One fantastic resource is Gary L. Gray's website at Penn State that includes everything LaTeX-related as it pertains to Macintosh: `http://www.esm.psu.edu/mac-tex/`. There are also three integrated packages for MacOS X (and two for classic MacOS) that simplify using LaTeX. The first is a relatively new project called TeXShop, which is only available for MacOS X. This program includes a text editor, and LaTeX compiler (it in fact uses the `tetex` distribution, although that is under the hood), and a document previewer based on `pdflatex`. TeXShop is thus an entirely integrated environment, and is totally Open Source. I suspect it will be ported to other versions of Unix in due time. You can download this package from: `http://www.uoregon.edu/$\sim$koch/texshop/`.

There are two shareware LaTeX packages that are available for both classic MacOS and MacOS X. Unlike TeXShop, these do not include integrated text editors (although they work well using Apple Events with the two text editors mentioned below). They do, however, compile .tex source files and display DVI files on the screen. They also have other useful features, including manipulating DVI files, creating postscript and PDF files from DVI files, and the like. The first is called OzTeX, which is available from `http://www.trevorrow.com/oztex/`. I have used OzTeX for nearly ten years, and find it to be a wonderful product, particularly for the meager price. Another option is CMacTeX, available from `http://www.kiffe.com/cmactex.html`.

Perhaps the best reason to use LaTeX on the Macintosh is the number of excellent text editors. Of course you can use Emacs or another Unix-based text editor in MacOS X (and, there are available ports for classic MacOS). But my time is spent using Alpha and BBEdit, which for my money are the two best text editors available on any platform. For pure LaTeX use I prefer alpha, which is available at `http://magnet.fsu.edu/$\sim$hall/docscripting/alpha/`. Alpha is a shareware program, and unfortunately is not yet available for MacOS X. Alpha is based on the tcl scripting language, and has a wonderful pallette of tools one can use to write LaTeX code. And, it integrates seamlessly with OzTeX and CMacTeX (a simple keystroke is all that is required to compile a document). BBEdit is a great multi-purpose text editor, particularly useful for HTML (`http://www.barebones.com/`). It too has some nice built-in LaTeX functions that make it as easy as possible.

In short, LaTeX is not only an extremely useful and powerful text processing system. It is an extremely useful and powerful *cross-platform* text processing system. For "the rest of us" – the 5% who do not use Microsoft Windows – LaTeX is a promising and viable text processing solution.

Review of Ron C. Mittelhammer, George G. Judge, and Douglas Miller's *Econometric Foundations*

Wendy Tam Cho
University of Illinois at Urbana Champaign
wendycho@uiuc.edu

Econometric Foundations. by Ron C. Mittelhammer, George G. Judge, and Douglas J. Miller. (Cambridge University Press, New York, 2000; 756 pp; $64.95. ISBN 0-521-62394-4.)

Give a man a fish and he will eat for a day. Teach him how to fish and you will feed him for a lifetime. Teach him the philosophy behind fishing and he will learn not only to fish, but to hunt and to farm as well. The lesson here is clearly that as with anything in life, establishing a firm foundation is the key to enduring success. Other options, in comparison, are clearly fleeting victories. In this philosophical sense, *Econometric Foundations* by Mittelhammer, Judge, and Miller (MJM), is a refreshing and much welcomed departure from the vast collection of econometrics texts. Whereas the goal of many econometric texts is to provide one with a set of tools, the goal of MJM is to help the student understand the tools, by giving the student a firm foundation in statistical theory.

Their mode for achieving this goal is simple. They begin with the most basic of models. Then, with each passing chapter, MJM tinker with the specification and generalize the reasoning behind the model. The clear overall logic of the book is an innovation that students and analysts will find extremely helpful. The book is separated into ten parts. It begins with a philosophical section on information processing and recovery. The second chapter jumps into regression models. The third section transitions into extremum estimators and nonlinear and nonnormal regression models. Section 4 examines how to avoid the parametric likelihood formulation. Section 5 looks at generalized regression models. In Section 6, they make a foray into simultaneous equation probability models and general moment-based estimation and inference. Section 7 discusses the all-important question of model recovery (variable selection and conditioning and the problem of noise covariance matrix specification). Section 8 treats the topic of limited dependent variable models. Section 9 makes something of a break and moves to Bayesian estimation and inference (though with a regression focus). The book ends with an epilogue that visits many of the issues of computer simulation and resampling methods that arise in the text. Throughout, MJM focus on establishing a firm base, developing a conceptual and empirical understanding of basic econometric models and procedures that provide the roots or foundations for variations found in specialized books and journal articles.

To boot, MJM is a valuable learning resource on multiple dimensions. The textbook is the tried-and-true medium, with nice pedagogical devices such as an "Idea Checklist—Knowledge Guidelines" and often "Computer Exercises" at the end of the various chapters. In addition, they provide a CD-ROM that includes examples from the book (written in GAUSS). The examples are especially helpful for learning because they are set up to be used interactively. For those who do not feel up to the statistical sophistication level that MJM assumes, the CD-ROM also has a primer on probability theory, classical estimation and inference, and ill-posed problems. Finally, a copy of GAUSS Light from Aptech Software is included, along with a short introduction to GAUSS and the complete GAUSS mailing list from 1995–1999. If all that were not enough, the book also has a web site (`http://www.econometricfoundations.com`) where one can download additional materials. For instance, instructors may download a solution manual, free of charge. Students will find example guides (in PDF format) that have additional background details for examples in the book. In addition, updates are made available for various aspects of the book, including a special discount offer to upgrade from GAUSS Light to the full version of GAUSS.

MJM is an especially attractive text for social scientists because it focuses on the type of data that social scientists usually encounter (incomplete, noisy, partial, etc.). They focus on practical, real-world data analysis rather than assuming that one lives in the purely theoretical data world. For political scientists, the book is most welcome for several reasons. First, it develops information processing and recovery from a viewpoint that is particularly amenable to the types of problems that we usually encounter. Second, it focuses on semi-parametric data based formulations, an especially useful but undertrodden path for us. Lastly, the electronic chapter on ill-posed inverse problems has direct application to many problems that are of particular interest to political scientists. For instance, the ecological inference problem is a classic case of an ill-posed inverse problem. There is no "solution" here, but the basis for how to think of the problem could not be more clear.

In short, MJM has all the virtues of a serious intermediate-level econometrics text, laudable foundational depth, numerous technological lagniappes in the form of the CD-ROM and the web site, and an expansive pedagogical vision. In the econometrics textbook market, this entrant is simply not duplicated, with its appeal spanning both the theoretical and the applied dimensions of statistical analysis. Comprehensive, up-to-date, innovative. Go fish!

Review of David Salsburg's The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century

Simon Jackman
Stanford University
jackman@stanford.edu

The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century. by David Salsburg (W.H. Freeman and Co, New York, 2001; 340pp; $23.95. ISBN: 0-716-74106-7.)

Twentieth century statistics is yet to get the detailed historical treatment accorded earlier periods of statistical development. Stigler's magisterial work, *The History of Statistics*, covers the emergence of modern probability up to the foundation of statistics as we would recognize it today. The pre-20th century period of statistical development is in some ways more interesting for political scientists than 20th century statistics; it is in the 18th and 19th centuries that political science (really, political economy) and statistics were at their intellectual perigee, the rise of the modern, secular, democratic state coinciding with "The Rise of Statistical Thinking" (the title of Porter's book on the subject), "The Taming of Chance" (Hacking) or "The Politics of Large Numbers" (Desrosières).

But for *political methodologists*, the last hundred years is where the interesting technical development takes place, where the foundations of workaday quantitative social science are laid down. It is in the 20th century that statistics emerges as a discipline in its own right, independent of its forebears, mathematics and political economy. Karl Pearson (1857-1936) is probably the last and most important tie with this earlier period; it is worth noting that Pearson's PhD was in Political Science from Heidelberg, and that he is probably the last major statistician whose primary training was not in statistics or mathematics. As Karl Pearson's generation of statisticians pass, they are replaced by giants such as R.A. Fisher, Jerzy Neyman, and Harald Cramér. And it is with this transition that statistics takes on it modern appearance, and where Salsburg's book begins.

Salsburg's title, *The Lady Tasting Tea*, refers to an incident where Fisher designed an experiment to assess whether one could really taste the difference between tea with milk added, or milk with tea added. Fisher is Salsburg's chief protagonist in his account of 20th century statistics, dominating the first half of the book and looming large over the second half. Salsburg is at his best in this early going, providing a well structured account of Fisher's contributions and intellectual trajectory. In particular, the story of Fisher's struggle with Karl Pearson makes for fascinating reading. According to Salsburg, Pearson delayed publication of Fisher's paper deriving the sampling distribution of the correlation coefficient in *Biometrika*, until Pearson's computational lab had generated tables for the distribution, with Fisher's analytics appearing as a footnote. Fisher then took up a position at the Rothamsted Agricultural Experimental Station, where his genius flourished and his influence on statistics as we know it today began to emerge.

The second half of the book is less well organized. Salsburg's approach to the subject matter is largely chronological and centered on particular persons or problems, dealt with in a series of short chapters (e.g., A.N. Komolgorov in "The Mozart of Mathematics", Frank Wilcoxon in "Doing Away with Parameters", assessing treatment effects in "Intent to Treat"; there are 29 chapters in all) rather than more broadly thematic. In addition, Salsburg's exposition is free of notation, equation, or even graphs. This makes for some jumbling of topics in the second half of the book; for instance, advances in Bayesian statistics spread out over different parts of the book. To the extent there is a hero in the second half of the book, it is John Tukey, dubbed the "Picasso of Statistics" by Salsburg. Statisticians such as George Box, W.E. Deming, Persi Diaconis, Brad Efron, and Don Rubin also garner chapters or several pages. It should also be noted that Salsburg stresses the contribution of women in a discipline almost totally dominated by men: Yvonne Bishop, Florence Nightingale David, Gertrude Cox, and Grace Wahba are singled out by Salsburg.

But all in all, there is a lack of coherence to this account of post-Fisher 20th century statistics, at least post-Fisher, which may reflect something about the subject matter itself, but, more likely, is grounded in Salsburg's choice about the level and tone of his exposition. My own preferences would have been to organize the material around the major themes of post-Fisher statistics. A common thread is the ongoing revolution in computation, turning theory into practice in the many areas: e.g., robustness, non-parametrics, visualization, the analysis of large data sets, and vastly simplifying and popularizing Bayesian approaches. Also missing is a treatment of econometrics or psychometrics. None of the stars of Mary Morgan's *The History of Econometric Ideas* gain a mention (e.g., Haavelmo, Tinbergen). Thurstone and Spearman do not gain mentions, nor does factor analysis. Intelligence testing via standardized tests is relegated to a brief discussion of the long list of applied problems Sam Wilks worked on, and the emergence of public opinion polling is considered an offshoot of large scale sampling by government agencies for generating official statistics.

Yet there is much we might take away from *The Lady Tasting Tea* for our teaching. Salsburg's treatment of Fisher tops my list in this regard. The strengths and limitations of Fisher's contributions are handled in a way that is extremely accessible. Salsburg reminds us of the colossal impact Fisher had: for instance, experimental design, ANOVA, maximum likelihood, the frequentist notion that sample statistics are random (estimators of fixed population parameters), and the characterization of properties of estimators (consistency, unbiasedness, efficiency), we owe to Fisher. Salsburg also provides a very even-handed and useful summary of the contributions of Jerzy Neyman and Egon Pearson in attempting to put statistical inference on a solid footing. The strengths and limitations of the Neyman-Pearson approach (one of the major points of contention between frequentists and Bayesians) are laid out quite clearly over four chapters in the middle of the book (although methodologists will want a more rigorous treatment, as in Barnett's *Comparative Statistical Inference* or Howson and Urbach's *Scientific Reasoning*). And sprinkled throughout the book are gems of insight, from interviews with famous statisticians. My favorite comes from Florence Nightingale David. She notes similarities between helping archaeologists dig for a kitchen midden and helping locate the launch sites of rockets targeting London remarks during WW2 and remarks: "It's curious there's a sort of unity among [statistical] problems, don't you think? There's only about half a dozen that are really different."

Salsburg's book is thus more than a collection of stories about twentieth century statisticians, the problems they worked on that made them famous, and the implications of particular advances. His book falls short of being a serious history of statistics, but is more than a compendium of statisticial biographies (e.g., Heyde and Seneta (eds), *Statisticians of the Centuries*). Salsburg is clearly writing for a broader audience than statisticians (or political methodologists). However, this is hardly the kind of book that will help our non-academic friends and family better understand what it is we do. In many ways, political methodologists are part of the target audience: not statisticians themselves, but more than interested by-standers. Political methodologists will be interested in the "behind-the-scenes" details Salsburg has used to tell the stories of twentieth century statisticians: the insights into the personalities, rivalries and conflicts, intellectual and professional struggles, choices made here, breakthroughs made there, and unresolved dilemmas. In short, *The Lady Tasting Tea* is light and enjoyable reading for those of us already familiar with contemporary statistics.

Review of Stephen M. Stigler's *Statistics on the Table: The History of Statistical Concepts and Methods*

J. Tobin Grant
Southern Illinois University
grant@siu.edu

Statistics on the Table: The History of Statistical Concepts and Methods. by Stephen M. Stigler (Harvard University Press, Cambridge, 1999; 448pp; $47.50. ISBN: 0-674-83601-4.)

When political methodologists recount the history of our subfield, it is likely to be brief. Most would likely begin our story in the 1960s when the use of statistics in political science became much more prevalent and political scientists began giving serious consideration to the use (and misuse) of statistics. But political methodology is really two stories. There is the familiar story of the subfield in political science, and there is also the less-known story of the rise of statistics to answer social questions. This is unfortunate, for as Stephen Stigler discusses in his excellent collection of essays on the history of social and behavioral statistics, *Statistics on the Table*, many of the bedrock concepts in statistics came from the need to apply rigorous statistical methods to answer important political questions.

The title of the book comes from a charge by Karl Pearson during a public debate over the effect of parent alcoholism on children. This debate took the form of newspaper articles between Pearson and such notable intellectuals as John Mayard Keynes. At issue was the effect of alcoholism on children—a nature vs. nurture debate in which Pearson, an advocate for eugenics, took the side that it was not alcoholism but the natural inferiority of the parents that affected children's educational achievement and health. While we would all quickly leap to argue against eugenics, we would face the same charge that Pearson gave to his detractors if we did so without doing our own analysis: "I am too familiar with the manner in which actual data are met with the suggestion that other data, if they were collected, might show something else to believe it to have any value as an argument. 'Statistics on the table, please,' can be my sole reply." In other words, Pearson was making the right claim that arguments about relationships in society must be based on empirical observations and subjected to statistical analysis. As the debate continued, the public was able to see some of the greatest minds in England debate questions that we still face today in our own research. What is the best way to measure behavior? How does one deal with sample selection? Under what conditions can a sample allow us to infer to a population?

Of the 22 chapters in *Statistics on the Table*, many, including the Pearson chapter, will prove to be of interest to those of us familiar with statistical concepts. To these readers I strongly recommend the book. The book will also be useful to our students who often encounter statistics late in their academic careers and are often more comfortable with narrative accounts than mathematical proofs. Selected chapters from this book will be a useful supplement to traditional textbooks, providing the instructor with a few more examples than the proverbial balls in an urn and Poisson's counting of kicks to the head during the Franco-Prussian war. That said, Stigler is a historian and as such finds some topics of interest that the average political methodologist would not. Chapters on obscure texts, the early uses of statistics in England, statisticians with little influence on modern techniques, and the measurements of standards are unlikely to be given more than a cursory read, but in all the book should be well-received by political methodologists.

The reason for this positive assessment is that Stigler does two things that all academics should strive to do. First, he writes well. His writing reveals his personality while accurately presenting the results of his rigorous research. Second, he asks interesting questions, questions that our students would likely ask about our material. Why is the normal curve called the "normal" curve? What is the difference between regression and the method of ordinary least squares, and why is it called regression anyway?

Statistics on the Table also spends a good deal of time on the question of eponyms-the practice of naming certain findings for people. As Stigler argues, very rarely are these the eponyms given to the original discoverer. Some of these mistakes are well-documented, including Pascal's triangle, the Cauchy distribution, and Cheychev's inequality, but others are not. Stigler writes several essays on discoveries such as the Gaussian distribution, Bayes Theorem (interestingly Stigler uses Bayesian methods to determine his answer), and the maximum likelihood. Again, many of these are useful introductions to these topics because they provide a narrative with which students can better understand the topic.

There is one chapter from *Statistics on the Table* that deserves particular attention from political methodologists. In his chapter on "Statistical Concepts in Psychology," Stigler asks why statistics began to be used in astronomy and other natural sciences by the 1820s, was employed by psychology in the 1860s, but was not used in social science until the 1890s. To give away one answer from the book—the goals of the different sciences are different and as such their use of statistics must be different as well. In astronomy the goal was to measure something true (based on Newtonian physics), but all observations contained some error. Statistical methods helped to discern the true parameters. In psychology such truth based on deductive theory did not exist, but with the use of experimental designs it became possible to have at least a baseline from which to compare. This comparison, however, demanded that psychologists use statistics to see if the differences between experimental groups were real or based simply on chance. For social scientists the task was much different, for social scientists "the statistical model itself defined the object of inference." Through statistical modeling, the social scientist was able to do what could not be done previously. Stigler concludes that "the role of statistics in social science is thus fundamentally different from its role in much of the physical sciences, in that it creates and defines the objects of study more directly. Those objects are no less real than those of physical science. They are even often much better understood." (199). It is easy to see how this goal of inference is true in political methodology, where much of our efforts are aimed at creating better estimators for modeling political phenomena.

As political methodologists continue pursuing this goal, they would do well to remember their roots. Our history is not simply the rise of statistics in political science. The history of political methodology must include the history of statistics, which Stigler finds is "broad in scope and rich in diversity, occasionally technical and complicated in structure, and never covered completely."

Using the Right Tools for Time Series Data Analysis

Patrick Brandt[1]
University of North Texas
brandt@unt.edu

Abstract

This brief note outlines some of the reasons for considering time series tools in data analysis software. A list of basic reasons for using time series software are presented and evaluated. An example using daily data from Goldstein et. al. (2001) is presented to illustrate how not using time series software to build time series data can lead to incorrectly dated data. Readers are urged to consider learning at least one major time series software package as part of their methodological toolkit.

Introduction

Time series data is defined as data that has some unit of time as the measurement unit. This important organizing property of the data means that there are unique functions and operations that are used in time series analysis. These include subsetting a sample based on dates, forecasting based on dates, using lagged values, differencing operations, and specific diagnostic tests based on time.

In general, many of these functions and operations are now included in even the most routine data analysis software. Almost every statistical software package can be coerced into making lagged variables (and thus taking $n'th$ differences), forecasting data, and subsetting data. That said, failing to use *time series software* for these functions can be a major mistake and make your work much harder. It may be a mistake because your software package *du jour* may not adequately handle the temporal properties of the data and not provide the correct functions and diagnostic tests needed to properly model time series data.

There are a number of qualities of time series data that lead us to want to use special software for its analysis. In large part, this is so we can utilize the temporal organization of the data in the analysis. Here I review some of the functions that are necessary for the analysis of time series data and why time series software (e.g., RATS, E-views, S-Plus, R, Ox, TSP) implements these functions better and more correctly than software that does not specialize in time series analysis (e.g., Stata or SPSS). The goal is not to advocate for any one software package or program.

Time Series Software Desiderata

What then are the benefits of time series software such as E-Views, Ox, RATS, R, S-plus, TSP, etc.? In part, this depends on what we want to do with time series data. We want to do different things with time series data than with cross-sectional data. These different needs mean that we need software that can handle these functions – and do it easily.

The things we would like to do with time series data are,

- Build, organize and subset data based on time.
- Graph and plot data based on time.
- Pre-tests, identification, and estimation for building autoregressive integrated moving average (ARIMA) models and diagnostics and specification tests for dynamic regression models.
- Construct advanced time series models (e.g., VAR, BVAR, SVAR, GARCH, Kalman filters, VECM, and error correction models).

Organizing and building time series

The first of these functions, building or subsetting data is common in all statistical software. However, for time series data we often want to set a sample and be able to revise it in the course of analysis. In many packages for cross-sectional data, the only way to do this will be to either "drop" observations from a sample or include sample selection statements with each estimation command. This complicates the analysis when all we want to do is rerun the same analysis with a new sample period.

A recent example presented exactly why time series software can provide an important check on the organization of data. As part of work for Brandt and Freeman (2002) I replicated the data series used in Goldstein et. al. (2001). The data consisted of Goldstein conflict/cooperation scores for 24 Middle Eastern country dyads based on Kansas Event Data System data. The data were based on daily counts of events that were then scaled and aggregated for the period from April 15, 1979 to June 30, 1997, a total of 6632 days over 18 years.

In the course of replicating the data using Phil Schrodt's Java KEDS Count program to aggregate events data for each day, I discovered that the Java KEDS Count program produced daily Goldstein scale series that were *not* highly correlated with those provided by Jon Pevehouse or a separate series provided by Phil Schrodt. The data provided by Schrodt was created using a Pascal version of the KEDS Count program.[2]

In an effort to pin down the source of the low correlations, I generated two datasets for comparison with the data that Schrodt provided from his Pascal KEDS Count program. I refer to this data provided by Schrodt as "Schrodt data." The first dataset I generated used the Java code version of the KEDS Count program to create an event series for 15 of the dyads.[3] I refer to this as "Java KEDS Count" data. The second dataset I created was constructed using Perl (to recode dyads) and Stata to aggregate and temporally organize the data. I refer to this as "Brandt" data. It should be noted that the original Pevehouse and Goldstein data used in Goldstein et. al. (2001) was created using an early version of the Pascal KEDS Count program (personal communication, Jon Pevehouse).

I then compared the daily counts of events in these 15 dyads (the Java KEDS Count and Brandt data) to those provided by Phil Schrodt (the Schrodt data). The results of the event correlations for the 6632 daily observations for 15 of the 24 dyads in Goldstein et. al. article are presented in Table 1.[4]

The immediate fact to note is the low correlations between the data produced by the Pascal version of the code (Schrodt data) and the other sources of aggregated events data for these dyads.

Further analysis revealed that the major discrepancies between the version of the code used by Schrodt and the Java version from the KEDS website were in the handling of leap days. From 1979-1997 there were 5 leap days (1980, 1984, 1988, 1992, 1996). All datasets, produced by each software program had the correct number of observations – the datasets had different values only in leap years. In each leap year, the leap days were handled incorrectly and led the time series to be shifted by one day using the KEDS Pascal code. The Java code incorrectly handled the leap days as well, entering a zero count for each dyad on the leap day. This was a minor error compared to shifting the days. However, it means that in six years, or one-third of the data, the incorrect values were entered for the days. This accounts for the correlations of 0.8 and lower for the series built using the incorrect date calendars in the Pascal KEDS Count software. The Brandt data made using Perl (to recode dyads) and Stata time series functions to aggregate the events into daily time series created a correct version of the data because Stata knew how to handle the leap days. A later version of the Schrodt and KEDS Count data with hand corrected dates to properly match the data to the leap days produced correct correlations (between 0.9 and 1).[5]

This error in handling leap days does not affect KEDS data aggregated with Schrodt's software for weekly or monthly aggregations. But it highlights the main point: using time series software to manage and organize time series data can be an important check on organization of the data. It does however affect the analysis of Goldstein et. al. (2001), since their Levant and Gulf event data was aggregated using a version of the Pascal code. Further work will be necessary to determine the substantive impacts of this data misalignment for the results reported in Goldstein et. al. (2001),

Dyad	Schrodt v. Java KEDS	Schrodt v. Brandt	Brandt v. Java KEDS
Iraq→USA	0.84	0.84	1.00
Israel→Palestine	0.81	0.81	1.00
Israel→Syria	0.81	0.81	1.00
Israel→Egypt	0.84	0.84	1.00
Israel→USA	0.77	0.77	1.00
Palestine→Israel	0.82	0.82	1.00
Palestine→USA	0.77	0.77	1.00
Syria→Israel	0.74	0.74	1.00
Syria→USA	0.79	0.79	1.00
Egypt→Israel	0.80	0.80	1.00
Egypt→USA	0.78	0.78	1.00
USA→Israel	0.78	0.78	1.00
USA→Palestine	0.79	0.79	1.00
USA→Syria	0.83	0.83	1.00
USA→Egypt	0.78	0.78	1.00

Table 1: Correlations in Event Series for Daily Middle Eastern Dyads based on KEDS data.

[2]All data used in this analysis and the Java KEDS Count program discussed here can be obtained from `http://www.ukans.edu/$\sim$keds`. In addition, Dale Thomas has written a new program to aggregate KEDS and TABARI data tha correctly handles leap days. It can be found at `http://www.ukans.edu/$\sim$keds/Thomas.Aggregation.zip`

[3]I focused on these 15 rather than the full 24 dyads used in the original Goldstein et. al. (2001) analysis because they were the data we received from John Pevehouse.

[4]These comparisons were done in R, to serve as a further check on the organization and coding of the data (and because R has time series functions and the facilities to read both raw text data and Stata files). The data and code can be obtained upon request.

[5]Minor discrepancies still remained in the datasets. These are due to differences in the coding dictionaries used by Schrodt and Pevehouse and Goldstein, as well as complexities in coding actors for Syria and Lebanon (personal correspondence, Phil Schrodt).

That said, one must be careful how one accounts for dates, even in time series software. For example, consider the following snippets of RATS code to set a daily calendar for the daily KEDS data discussed above:

```
CALENDAR 1979 1 365
```

and

```
CALENDAR(SEVENDAY) 1979 1 1
```

In the first version, the '1' tells RATS that the data are annual, with 365 subperiods per year. This tells RATS that every 365 observations the year counter should increase by one unit. It does not, however, invoke the internal perpetual date functions to account for the days. The latter version does this. It uses the perpetual date functions to account for the actual days in seven day weeks (as opposed to 5 day weeks for financial time series). Instead of saying we have a year that is subdivided, this latter formulation of the calendar uses days as the unit of analysis, not a subdivided year. This means that leap days are recognized, dates can be aggregated properly into months and quarters, and data can be subsetted by day, week, and month.[6]

This last example highlights another important issue with time series data – accounting for the frequency and periodicity of the data. It is useful for the software to be able to determine whether the series is annual, quarterly, monthly, etc. This means that searching for seasonal and cyclical patterns can be more easily accomplished. Such capabilities are more important as high frequency data are employed in more and more political science applications such as Goldstein et. al. (2001), Brooks, Hinich, and Moiyneux (2000), Herron (2000), and Herron et. al. (1999).

Graphing and plotting time series data

Graphing time series data would appear simple. But a little experience shows that this is not always the case.[7] In non-time series oriented software it can be more complicated, particularly when the user knows the starting and ending dates of the data series, but as is typically the case they are not recorded in the dataset with a date variable. Consider graphing such a time series in Stata. To graph a time series variable `y`:

```
tsset y         /* sets the date functions */
                /* on for data             */
gen constant=1  /* creates a vector of 1's */
gen t = sum(c)  /* turns the vector of 1's */
                /* into a counter          */
                /* or trend                */
graph y t       /* Graph the series        */
```

The added complexity here is the need to either make or use another time series variable, here `t` to act as the x-axis in the plot – even after we have told Stata that the variable is a time series.[8] In addition, the time scale of the variable `y` is not accounted for as an intrinsic attribute of the series. It must be supplied. This introduces an additional source of error, since the user must supply the appropriate time variable, which may be different for series of different lengths in the same dataset due to missing values, lags, or other transformations. So, while Stata includes a number of time series analysis functions, its representation of time series data must be supplied by the user as part of that data.

Other time series software handles this task very simply. For example, using the time series functions in R/S-Plus:

```
y.timeseries<-ts(y,start=c(0,1),freq=1)
plot(y.timeseries)
```

or in RATS

```
GRAPH 1
# Y start end
```

Both of which produce a time plot over the indicated sample. Subsetting the data can be done on the fly as part of the graphing functions (using the `window()` function in R, or setting start and end times in RATS), and additional series or forecasts can easily be added to the data. Note that we can set both the periodicity and frequency of the data and use it to act as a property of the data for plotting. Further, the information about the time series properties of the variables is an intrinsic attribute of the data and is recognized by the program when plotting the data. So while Stata works well for organizing time series data (as seen in the last section), it takes four lines of code to do what RATS and R can do in two lines. This difference can start to add up if we need to make multiple plots over different time horizons.

[6] Thanks to Tim Hellwig for this example and verifying the implications of these two approaches with Estima.

[7] In fact, it was the need to graph arrays of Monte Carlo results from time series models that forced me to learn S-Plus and R.

[8] There are Stata functions available from various web archives that allow users to plot time series data without doing this. For the occasional user this adds more complexity since they have to know where to find and how to install and use the new function. The user still must use a separate time series variable to set the calendar for the data with the `tsset` function. See for example the Stata function `sssplot` at `http://ideas.uqam.ca/ideas/data/Softwares/bocbocodeS329601.html`.

ARIMA modeling and dynamic regression

A third reason that recommends the use of time series software is the tests and estimators used for standard time series analysis. By this I mainly mean Box-Jenkins ARIMA specification and modeling. These functions include autocorrelation functions, partial autocorrelation functions, Portmanteau tests, and unit root tests.

The application of these models is greatly simplified in standard time series programs. A good example of this are unit root tests, such as the commonly used augmented Dickey-Fuller tests. These can all be implemented by simple regressions on lagged variables, differenced variables, and a time trend (see Hamilton (1994) for the details of the implementation of these tests and the various non-standard critical values). However, for comparisons of the tests across different specifications (of the time trend and augmentation) one must use comparable sample periods. This is not typically handled well by non-time series software, because it is harder to track the dates of observations. For the tests to be comparable, they should be computed over identical sample periods – a task made much easier in time series software where the start and end times for a sample can be set directly for each estimation command. This is particularly critical in small samples.

Another example though illustrates where time series software can model data in more complex ways and do so much more easily than non-time series software. Consider the following time series intervention model with compound transfer functions for monthly data:

$$y_t = \frac{(\omega_0 + \omega_1 L)}{1 - \theta_{12}^X L^{12}} X_t + \frac{(\eta_0 + \eta_1 L)}{1 - \theta_{12}^Z L^{12}} Z_t + \frac{1}{(1 - \phi_1 L - \phi_2 L^2 - L^{12})} u_t$$

where L is the lag operator and the superscripts on θ keep track of which denominator the lag effect comes from. This is an ARIMA$(2,0,0)(3,1,2)_{12}$ model with two compound interventions. This example was inspired by an illustration of intervention models with John Williams for a time series course. The example involved data on monthly tons of trash collection in Bloomington, Indiana. Two permanent policy interventions, the institution of a tag system for each can of trash and the banning of yard waste such as grass clippings, were fit with this model. The complex dynamics of the data arose because of the seasonal components of the trash collection and the changes in the dynamics at each of the intervention points.[9] Such a model can be easily estimated in RATS. For example in RATS we might use code such as:

[9] The seasonal non-stationarity appeared to be due to the surge in trash created when the students moved out of their apartments in May of each year.

```
boxjenk(ar=2,sdiffs=1,sar=3,sma=2,inputs=2) y
# x 1 1
# z 1 1
```

We can directly estimate the quantities of interest for this transfer function model with this RATS code and obtain the intervention impacts and the polynomial coefficients. The output will also produce standard errors for these parameters. In addition, we will be estimating only the eight parameters we need for the model.

However, in Stata we have to expand out the polynomials and estimate the non-consecutive lags:

```
arima y x L.x L2.x L3.x L12.x L13.x L14.x
        L15.x L24.x L25.x z L.z L2.z L3.z
        L12.z L13.z L14.z L15.z L24.z L25.z,
        ar(1 2 12 13 14 24 25 26 36) ma(12 24);
```

The results of this estimation will not return the quantities of interest for the transfer function model. We will have to compute them by equating the coefficients with the polynomial coefficients for the expanded model. Once we have solved back for the intervention effects and the polynomial coefficients for the compound transfer function model, we will not have standard errors. Further, this model will estimate 31 parameters, using nearly four times as many degrees of freedom as the RATS model! The other, less desirable alternative, is to code the likelihood function for the model in Stata and use the optimization routines in `ml` to compute the estimates.

Advanced Time Series Models

Finally, advanced time series models are generally only included in time series software. Estimation of reduced form identified VAR models can be done easily in any regression package using equation-by-equation estimation. However, the inversion of the VAR to produce impulse responses requires complex matrix computations. Monte Carlo integration to produce error bands for the impulse responses is still more complex.

In addition, more specialized models such as GARCH, Kalman filters, vector error correction models (VECM), and Markov switching models are typically only available in time series programs. Some of these can be easily estimated using standard regression software (such as VECM), however, the computation of the quantities of interest for these models can be complex and is best left to specialized software.

Using specialized time series software for new time series applications is a must since the programming to implement these techniques can be difficult. This includes some of the recent developments in Bayesian time series analysis such as those for Bayesian VAR conditional forecasting using Gibbs sampling (Waggoner & Zha 1999), likelihood based bands for VAR impulse responses (Sims & Zha 1999). Recently, I have programmed some of these methods using R/S-Plus. This would have been impossible without the specialized time series functions and programming capabilities of this software.[10] These Bayesian time series models are at the forefront of time series analysis and will see more use in political science as tools for evaluating forecast uncertainty, hypotheses and counterfactuals (e.g., Brandt and Freeman 2002). As Bayesian techniques have become more widespread in political science and are applied to time series problems, specialized time series software such as R, S-Plus, RATS and Ox that allow for user designed simulations and classes for Monte Carlo simulation are more appealing. They allow users to both program new methods and utilize existing time series techniques in a single software program.

Conclusion

Why then should one learn at least one time series software package? As the example above shows, it can save you considerable time in preparing time series data, since it can account for the basic organization of the data over time. Second, when one does need to diagnose and fit time series models, one will have the complete complement of univariate and multivariate time series techniques available in standard time series software.

More complex is the issue of "which time series software to use?" This essay has not been an attempt to review the major time series software packages on the market. Rather, it is an attempt to tell practitioners that they need to consider the organization of their time series data and how the data will be modeled. The choice of any one package is up to the analyst, but the need to use a specialized time series program or functions is important and should be given some thought when beginning a time series data analysis project.

References

Brandt, Patrick T. & John Freeman. 2002. "Moving Mountains: Bayesian Forecasting As Policy Evaluation." presented at the 2002 Meeting of the Midwest Political Science Association.

Brooks, Chris, Melvin J. Hinich & Robert E. Moiyneux. 2000. Episodic Nonlinear Event Detection Political Epochs in Exchange Rates. In *Political Complexity: Nonlinear Models of Politics*, ed. Diana Richards. Ann Arbor: University of Michigan Press pp. 83–98.

Goldstein, Joshua S., Jon C. Pevehouse, Deborah J. Gerner & Shibley Telhami. 2001. "Reciprocity, Triangularity, and Cooperation in the Middle East, 1979-1997." *Journal of Conflict Resolution* 45(5):594–620.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Herron, Michael C. 2000. "The Source of Gridlock: Divided Government or Preference Conflict." http://icg.harvard.edu/~gov3009/fall00/herron.pdf.

Herron, Michael C., James Lavin, Donald Cram & Jay Silver. 1999. "Measurement of Political Effects in the Stock Market: A Study of the 1992 Presidential Election." *Economics and Politics* 11(1):51–82.

Sims, Christopher & Tao Zha. 1999. "Error Bands for Impulse Responses." *Econometrica* 67(5):1113–1156.

Waggoner, Daniel F. & Tao Zha. 1999. "Conditional Forecasts in Dynamic Multivariate Models." *Review of Economics and Statistics* 81(4):639–651.

[10] In fact, I abandoned using Gauss because it lacked the time series capabilities of R.

The Society for Political Methodology

ad indicia spectate

Section Activities

Announcements from the President

Jonathan Nagler
New York University
Jonathan.Nagler@NYU.edu

Host the 2003 or 2004 Annual Summer Political Methodology Meetings

We are looking for sites to host the 2003 and 2004 Annual Summer Political Methodology Meetings. If you would like to bring the best and brightest of Political Methodologists to your campus for 3 to 4 days, please consider submitting a proposal to host the meeting.

Hosting the meeting is a great way to promote your department to the discipline, and to help give your graduate students exposure to a very high quality meeting. Each year the Summer Meeting brings about 100 participants together for 3 days of intensive interaction. And this generally includes the majority of people doing the very best work in quantitative political methodology. The meeting normally lasts from Thursday morning through Saturday afternoon or Sunday morning. The host institution takes responsibility for housing and feeding the participants. Housing has traditionally been provided on-campus in undergraduate dorms. Faculty participants are generally charged a 100−150 registration fee, which is used to help cover the expenses of the host institution. The Society for Political Methodology covers airfare for many participants through section funds and a grant from the NSF. As a courtesy most host institutions have provided participants with the option to stay in off-campus hotels, at the participants' expense.

You can find information about past meetings at the Political Methodology website: `http://web.polmeth.ufl.edu/conferences.html`.

If you are interested and would like additional information, please contact Jonathan Nagler, Chair, Summer Site-selection Committee. jonathan.nagler@nyu.edu; 212-992-9676.

The search for the successor editor for *Political Analysis* set to begin

Under Neal Beck's reign as editor the emergence of *Political Analysis* as a quarterly has been a great success. Neal will shortly step down as editor, and we will be searching for a successor. A formal search committee will be announced shortly. But if you have suggestions or a nomination, please send them to Jonathan Nagler (jonathan.nagler@nyu.edu); they will go to the search committee as soon as it is appointed.

Points of Interest in the 2002 ICPSR Summer Program

There have been some changes at ICPSR that may be of interest to you, your colleagues, and students. In particular, the ICPSR Summer Program will offer "The Enhanced Summer Program Cluster of Advanced Courses". The Enhanced Summer Program Cluster is an integrated cluster of three advanced courses designed to provide significantly enhanced statistical and computational skills. This new program, offered in the first four week session, will link the existing Bayesian and Maximum Likelihood courses, add a much needed course in "Modern Regression", and provide a common computing environment for all three based on the S statistical language. By the "S language" we mean to include both S-Plus and R implementations, but the program will stress the use of R for computing so that students can take the free R program home with them.

This new cluster of courses will provide a very integrated training experience–students will take a common course in S, will typically take two of the three advanced statistics courses, and will be pressed to make connections between the courses, rather than view them as independent. Because all three courses will share the S foundation, students will have a common computational environment, one that has become the de facto standard in statistics and which has a growing presence in the social sciences. This common language is more than just a shared computing interface–it is a common way to think about the structure of models, estimation, and display of data. Immersing students in this common culture will allow them to substantially increase their sophistication in both statistical modeling and computing during the four weeks of the first session.

The new course in Modern Regression will cover regression with much more emphasis on graphical displays, diagnostics, and non-parametric fits than does the current more "econometric" perspective. While the topic is "regression", this is a more advanced and sophisticated course than currently offered in the first session. It would be appropriate for anyone who has had a regression course and wants to review it, but the real target will be those who want to see what's new in regression and graphics.

The Maximum Likelihood course is revised to take advantage of the S language, better graphical methods for interpretation and will put more emphasis on Generalized Linear Models as a statistical framework. The Bayesian Inference course will make explicit links to Likelihood, and will continue its current use of the S language. The S language is particularly useful in this course because the standard entry-level implementation of Markov chain Monte Carlo (MCMC) is built upon S-like syntax (BUGS).

In addition, there is an integrating special seminar series in Advanced Statistical Computing that links the three statistics classes. It will provide not only an introduction to the S language, but will support all the advanced skills needed for the classes and especially beyond. The goal will be to achieve enough fluency in S that students will think it natural to continue using S when they return home from the Summer Program. The course will also introduce LaTeX for document creation, showing students how to incorporate math, text, tables and graphics in elegant documents.

The S language is not merely a powerful computational tool, but is a common language that can help students appreciate the linkages between these statistical topics. With a single tool that covers the range of applications this cluster represents, students will return home with dramatically improved analytic abilities. The dominance of the S language in statistics journals means that the provided training is in the use of a tool that is almost certain to be the basis for statistical applications for decades to come.

The set of classes will be taught by John Fox, McMaster University; Charles H. Franklin, University of Wisconsin, Madison; Jeff Gill, University of Florida; Hank Heitowit, University of Michigan; William Jacoby, University of South Carolina.

In addition, the Summer Program will again offer "Complex Systems Models in the Social Sciences," sometimes this area is referred to as "adaptive systems" or "agent-based models" (The type of modeling often identified with the Santa Fe Institute). Instructors are Ken Kollman and Scott Page, University of Michigan.

There are also one week courses on:

- Categorical Analysis: Introduction to Regression Models (instructor: Scott Long)
- Network Analysis (instructor: Stanley Wasserman)
- Mixed Models for Categorical Data
- "LISREL" Models (instructor: Ken Bollen)
- Spatial Analysis (instructor: Luc Anselin)
- Event History Analysis

Standard 4 week courses include:

- Maximum Likelihood Estimation (instructor: Charles Franklin)
- Advanced MLE
- Scaling & Dimensional Analysis (instructor: Bill Jacoby)
- Time Series (Instructor: John Williams)

Finally there is a 2 course sequence in formal modeling:

- Game Theory (instructor: Mark Fey, Rochester)
- Rational Choice Theories (instructor: Jim Johnson, Rochester)

These are just some of the highlights of the 2002 program. You can find the full Program Announcement and the on-line registration form in the Program web site: `www.icpsr.umich.edu/sumprog/`. If you have further questions, please contact: Henry Heitowit, Director Educational Resources, ICPSR, P.O. Box 1248, Ann Arbor, MI 48106-1248. `http://www.icpsr.umich.edu/sumprog/` voice: (734)998-9888 FAX: (734)998-9889.

Notes on the Essex Summer Program

This year's Essex Summer School in Social Science Data Analysis and Collection will offer over 50 one and two-week introductory, intermediate and advanced courses from July 6 through August 16 on topics which include: social survey design and analysis, sampling, regression, multilevel analysis, time series analysis, latent class analysis, discourse analysis, game theory, rational choice, social theory, data visualisation and data mining, social network analysis, structural equation models, logit, probit and other generalised models, maximum likelihood estimation and limited dependent variables, geographical information systems, socio-legal research methods, qualitative data analysis, focus groups, interviewing, participant observation, content analysis, SPSS, Amos, Stata, British Household Panel Survey, time budget collection and analysis and comparative policy analysis.

New courses this year include: Bayesian Methods for Social Science Data Analysis, Conflict Modelling and Analysis, Simultaneous Equation Models, Ecological Analysis, Multidimensional Scaling, Introduction to Stata, Scale Analysis: Developing Measurement Instruments.

A small number of ESRC bursaries are available to participants from British academic institutions.

For further details see `http://www.essex.ac.uk/methods` or e-mail sumsch@essex.ac.uk or write to The Essex Summer School in Social Science Data Analysis and Collection, University of Essex Colchester, Essex CO4 3SQ, United Kingdom or Fax [international] 44-1206-873598 [UK/Eire] 01206-873598 or telephone [international] 44-1206-872502 [UK/Eire] 01206-872502.

EITM Summer Training Institute Announcement

The National Science Foundation's Political Science Program is supporting four annual four-week summer institutes, to be held from 2002-2005 at Harvard, Michigan, Duke, and UCBerkeley, respectively. Each of these institutes will accommodate up to 25 advanced graduate students and junior faculty. Funding is available to defray the cost of participants' travel, accommodation, and subsistence. Programs will be selective. Admission will be based significantly on the quality and potential of research presented. A team of up to 15 research faculty will conduct each institute. Training offered will include teaching and research components, providing students a high degree of individualized interaction with a far wider and deeper array of mentors than is available at any individual institution. The first institute will take place at Harvard from June 24 - July 19, 2002.

The Empirical Implications of Theoretical Models (EITM) focus of the Political Science Program at the NSF recognizes that gaps have appeared between theory and empirical method, and that these gaps impair scientific progress. The scientific study of politics requires empirical evaluation of theoretical models, but theories are often produced without adequate empirical exploration and empirical work too frequently adopts sketchy and oversimplified theory. To ameliorate this we need to train a new generation of scholars who can better link theoretical and empirical work, by offering younger scholars an opportunity to learn by seeing and doing in conjunction with older scholars who have been leaders in advancing theoretical and empirical work, focusing on substantive areas where appreciable research integrating theory and methods already exists.

Areas of instruction will be drawn from among spatial models, institutional analysis, macro- and international political economy, bargaining and coalitions, and international security. Formal models will include game theory, differential equation dynamic models, simple decision theory, and more complicated behavioral decision-making models. The empirical toolkit will encompass not only statistical inference but also focused analytically-based case studies, experimental methods, and computational models.

Lead participants from the four sites, in the order they will host the institutes, are James Alt, Harvard University; Rob Franzese, University of Michigan; John Aldrich, Duke University; and Henry E. Brady, UC Berkeley. All will be involved in the 2002 program.

Applying

We intend to accept a pool mostly of advanced graduate students, who are past general examinations, preferably modally with a dissertation prospectus or plan in hand but not yet at the writing-up stage. We will also accept a few junior faculty, preferably again at a stage where they are either wishing to add something to the (completed) dissertation before publication, or embarking on a "second project". The pool will number about 25. Admission will be based on the quality and potential of research proposed in the application.

Interested candidates should apply by providing:

- a vita with name and contact information
- current location and position
- a 5-10 page description of a research proposal
- a short (1-2 page) statement of interest and purpose in applying for the summer program
- two letters of recommendation.
- if student: current status in graduate school (past exams? defended proposal?).

It is preferable but not mandatory that application materials, including the two letters of recommendation, be submitted as PDF attachments (alternately in Rich-Text-Format) via e-mail.

All application materials must be received by Friday May 10th at eitm@latte.harvard.edu. Or if necessary, by mail to:

EITM Coordinator CBRSS, Harvard University 34 Kirkland St. Cambridge, MA 02138

Successful applicants will be notified (by e-mail) no later than Friday, May 24th. They will receive a travel and living expense stipend. For out-of-town participants the stipend will cover round-trip airfare and dormitory housing. All participants will receive a modest allowance for meals and incidentals.

Design of Instruction

The scientific study of politics requires empirical testing of theoretical models, but theories are often produced without adequate testing and empirical work too frequently uses sketchy and oversimplified theory. Gaps have appeared between theory and empirical method, and these gaps impair scientific progress. The goal of the Empirical Implications of Theoretical Models (EITM) program is to train a new generation of scholars who can better link theory and empirical work.

We will concentrate on areas where research integrating both theory and methods already exists. There will be three units in 2002, detailed below. Each unit will explicate the steps needed to construct a "test" of a model such as by considering the validity of its basic assumptions or by developing conjectures from comparative statics and other deductions from the model. Once this is done, methods must be developed to see if the data confirm or reject the model. This could involve specifying test equations with the proper control variables and functional forms, deriving statistical estimators, designing an experiment, or framing a simulation.

The first three week-long sessions will be organized as follows. Morning and early afternoon sessions will be devoted to presentation of materials. A daily early evening session and Saturday sessions will be devoted to research presentations by faculty and students. Evenings will be lab time for exercises as necessary. Each teaching unit will also feature two guest lecturers whose role will be to present some completed pieces of research that exemplifies the integration of formal theory and empirical methods (e.g., theory in the morning, empirical application in the afternoon). The guest lecturers will appear after enough of the substance of the area has been covered to allow students to participate in critical appreciation and evaluation of the work presented. To combine integrative teaching of formal theory and empirical methods with research presentation and interaction, the fourth week of each session will provide an intense mentoring experience. A significant purpose of this unit will be to germinate new research ideas.

Contact: Alison Ney, Program Coordinator, Center for Basic Research in the Social Sciences, Harvard University, aney@latte.harvard.edu

Notes from the Editor of *Political Analysis*

Neal Beck
New York University
neal.beck@nyu.edu

Volume 10(1) has been mailed and (2) is in final stages of production for mailing in late April or May. (3) is a special issue on spatial and geographic methods, with most papers having come from a conference at the University of Colorado, Boulder, which brought together geographers and political scientists to discuss ways of bringing the two disciplines closer together. (4) is another special issue on experimental methods. 11(1) is already full, and (2) is filling up rapidly.

Volume 11 will be my last volume, a new editor is being sought to take over starting with Volume 12. It has been great fun editing Political Analysis and I would like to thank all the authors who have put up with my various requests and demands, and who have made Political Analysis a success (I think!).

Forthcoming in 10(3)

John O'laughlin, University of Colorado, "The Electoral Geography of Weimar Germany: Exploratory Spatial Data Analyses (ESDA) of Protestant Support for the Nazi Party".

Patrick Heagerty, University of Washington, Michael D. Ward, University of Washington and Kristian Skrede Gleditsch, UCSD, "Windows of Opportunity: Window Subseries Empirical Variance Estimators in International Relations".

Brady Baybeck, University of Missouri-St. Louis and Robert Huckfeldt, Indiana University, "Spatially Dispersed Ties Among Interdependent Citizens: Connecting Individuals and Aggregates".

Luc Anselin and Wendy K. Tam Cho, University of Illinois, "Spatial Effects and Ecological Inference".

Michael Ward, University of Washington and Kristian Gleditsch, UCSD, "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace".

10(4)

Rick Wilson, Rice University, "Fairness and Rejection in the Ultimatum Bargaining Game".

Rose McDermott, Cornell University, "Experimental Methodology in Political Science".

James L. Gibson, Washington University, "The Role of Theory in Experimental Design: Experiments without Randomization".

Howard Lavine, SUNY Stony Brook, Milton Lodge, SUNY Stony Brook, James Polichak, University of Michigan and Charles Taber, SUNY Stony Brook, "Explicating the Black Box through Experimentation: Studies of Individual Differences and Cognitive Processes".

Donald P. Green and Alan S. Gerber. Yale University, "The Downstream Benefits of Experimentation".

Adam F. Simon and Tracy Sulkin, University of Washington, "The Impact of Discussion on Political Decisions: An Experimental Approach".

Volume 11

Robert Voogt and Willem Saris, Amsterdam, "To Particpate or Not to Participate: Modeling Survey and Political Participation".

Wijbrandt H. van Schuur, Groningen, "Mokken scale analysis: a nonparametric probabilistic version of Guttman scaling for survey research".

Anne E. Sartori, Princeton University, "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions".

Michael C. Herron and Kenneth W. Shotts, Northwestern University, "Using Ecological Inference Point Estimates as Dependent Variables in Second Stage Linear Regressions".

Keith Krehbiel, Stanford University, "The Coefficient of Party Influence."

Robert P. Berrens, University of New Mexico, Alok K. Bohara, University of New Mexico, Hank Jenkins-Smith, Texas A&M University Carol Silva, Texas A&M University, and David L. Weimer, University of Wisconsin, "The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples".

R. Michael Alvarez, Robert P. Shermanz, Carla VanBeselaere, Caltech, "Subject Acquisition for Web-Based Surveys".

THE POLITICAL METHODOLOGIST
Department of Political Science
Pennsylvania State University
University Park, PA 16802

 I gratefully acknowledge the support of the Department of Political Science of the Pennsylvania State University in helping to defray the editorial and production costs of the newsletter.

Subscriptions to *TPM* are free to members of the APSA's Methodology Section. Please contact APSA (202 483-2512, `https://www.apsanet.org/about/membership-form-1.cfm`) to join the section. Dues are $25.00 per year and include a free subscription to *Political Analysis*, the quarterly journal of the section.

Submissions to *TPM* are always welcome. Articles should be sent to the editor by e-mail (sdeboef@la.psu.edu) if possible. Alternatively, submissions can be made on diskette as plain ascii files sent to Suzanna De Boef, Department of Political Science, 108 Burrowes Building, Pennsylvania State University, University Park, PA 16802. LaTeX format files are especially encouraged. See the *TPM* web-site [http://web.polmeth.ufl.edu/tpm.html] for the latest information and for downloadable versions of previous issues of *TPM*

TPM was produced using LaTeX on a PC running MikTex and WinEdt.