The Political Methodologist

Newsletter of the Political Methodology Section
American Political Science Association
Volume 6, Number 2, Spring 1995

Co-Editors: R. Michael Alvarez, California Institute of Technology
Nathaniel Beck, University of California, San Diego

Contents

Notes From the Co-Editors

R. Michael Alvarez
California Institute of Technology

Nathaniel Beck
University of California, San Diego

We begin this issue of *The Political Methodologist* with more discussion about the role of replication in empirical political research. In this issue we have an article by Jonathan Nagler which provides wonderful guidelines which we should strive to follow on how to write clear and understandable computer code, and thereby engage in better scientific practice. Many of the problems people commonly run into when trying to reproduce their own empirical work, and especially the work of others, could be resolved if the guidelines suggested by Nagler are followed.

With this issue of *The Political Methodologist* we also hope to stimulate discussion about how political methodology is taught in our graduate programs. We begin with a symposium on an important book on methods, *Designing Social Inquiry*. Those who were fortunate enough to be in the audience at the last APSA meeting for the lively commentary and discussion on this book already know that the commentary by Larry Bartels and Henry Brady mixed wit, criticism, and praise. We print the expanded versions of their commentary here to provide a wider forum for the discussion of qualitative methodologies in political research. Then, we have two discussions of the graduate methodology programs at Berkeley and Rochester. We also include a significant syllabus on research methods in international relations.

Another subject of discussion recently has been the role of formal models in political science. Here we include an essay by Scott Gates and Sherry Bennett Quiñones on the integration of formal and empirical political research agendas.

Last we have received a number of books which might be of interest to political methodologists, some of which are reviewed below. These range from books which many of you might find useful for your own libraries to books which should be thought of as candidates for courses in political methodology.

Coding Style and Good Computing Practices

Jonathan Nagler
University of California, Riverside

Recently *The Political Methodologist* presented the case for replication.[1] Replication depends upon individual researchers being able to explain exactly what they have done. And being able to explain exactly what one has done requires keeping good records of it. This article describes basic good computing practices. It contains advice for writing clear code, as well as advice on general computing practices to maintain. The goals are simple. First, the researcher should be able to replicate his or her own work six hours later, six months later, and even six years later. Second, others should be able to look at the code and understand what was being done (and preferably why it was being done). In addition, following good computing practices encourages the researcher to maintain a thorough grasp of what is being done with the data, and thus makes it easier to perform additional analyses. Good coding allows for more efficient research. One is not always re-reading one's own work and retracing one's own steps to perform the smallest bit of additional analysis.

The sequence this article is written in leads the reader from more general themes to more specific ones. I encourage readers who don't have the patience for the big picture to skip ahead rather than skip the whole article. Even learning basic conventions about variable names will put you ahead of most coders! And this article is *not* meant only for sophisticated statistical researchers. In fact the statistical procedures you will ultimately use has absolutely nothing to do with the topic of this article. These practices should be used even if you are doing nothing more complex than producing 2x2 with a particular data-set.

First, what do I mean by computing practices and 'code'? Computing practices covers everything you do from the time you open the codebook to a dataset, or begin to enter your own data, to the time you produce the actual numbers that will be placed in a table of an article. By code I mean the actual computer syntax -- or computer program -- used to perform the computations. This most likely means the set of commands issued to a higher-level statistics package such as SAS or SPSS. I will refer to a given file of commands as a 'program.' Most political scientists do not think of themselves as computer programers. But when you write a line of syntax in SAS or SPSS, that is exactly what you are doing -- programming a computer. It is coincidental that the language you use is SAS instead of Fortran or C. The paradox is that most political scientists are not trained as computer programmers. And so they never learned basic elements of programming style, nor are they practiced in the art of writing clean, readable, maintainable code. The classic text on programming style remains Kernighan and Plauger *The Elements of Programing Style* (1974), and most books on programming include a chapter on style. I recommend including a section on coding style in every graduate methods sequence.

This article starts from the point when a raw data set exists somewhere on a computer. It breaks analysis down into two basic parts: coding the data to put it into a usable form, and computing with the data. The first part can be broken down into two component steps: reading the essential data from a larger data-set; and recoding it and computing new variables to be used in data analysis.

Here are the basic points that will be covered below, and at the end of the article I list a set of 'rules' laid out.

- Labbooks: Essential.
- Command Files: they should be kept.
- Data-Manipulation vs. Data-Analysis: these should be in distinct files.
- Keep Tasks Compartmentalized ('modularity').
- Know what the code is supposed to do *before* you start.
- Don't be Too Clever.
- Variable Names should *mean* something.
- Use parentheses and white-space to make code readable.
- Documentation: All code should include comments meaningful to others.

Labbooks

Our peers in the laboratory sciences maintain lab-books. These books indicate what they did to perform their experiments. We are not generally performing experiments. But we have identical goals. We want to be able to retrieve a record of every step we have followed to produce a particular result. This means the lab-book should include the names of every file of commands you wrote, with a brief synopsis of what the file was intended to do. The labbook should indicate which produced results worth noting. For instance, the following labbook entry describes what the file of Gauss code 'mnp152.g' did:

```
Date: Oct 11, 1994
File: mnp152.g
Author: JN
Purpose: Analysis - this file does
         multinomial probits on
         our basic model.
```

[1] I thank Charles Franklin, Bob Hanneman, Gary King, and Burt Kritzer for useful comments and suggestions.

```
Results: The results were used in
         Table 1 of the Midwest paper.
Machine: Run on billandal (IBM/RS6000).
```

It is a good idea to have a template that you follow for each Labbook entry, this encourages you to avoid getting careless in your entries. The above template includes: Date, File, Author, Purpose, Results, and Machine. You might have a set of 'Purposes' – – Re-Coding, Data-Extraction, Data-Analysis – – that you feel each file fits into. It may seem superfluous to indicate what machine the file was executed on. But should you develop the habit of computing on several machines, or should you move from one machine to another in the course of a project, this information becomes invaluable in making sure you can locate all of your files.

It makes a lot of sense to have the labbook online. It can be in either a Wordperfect file, or a plain ascii text file, or whatever you are most comfortable writing in. First, it is easy to search for particular events if you remember names of them. Second, it can be accessed by more than one researcher if you are doing joint-work.

> **Rule:** Maintain a labbook from the beginning of a project to the end.

Labbooks should provide an 'audit-trail' of any of your results. The labbook should contain all the information necessary to take a given piece of data-analysis, and trace back where all of the data came-from - including its original source and all recode steps. So while there are many different sorts of entries you might want to keep in a labbook, and many different styles to keep them; the central point to keep in mind is that whatever style you choose meets this purpose.

Command Files

The notion of 'command files' may appear as an anachronism to some. After all, can't we all just work interactively by pointing and clicking to do our analysis - not having to go through the tedium of typing separate lines for each regression we want to run and submitting our job for analysis? Yes, we can. But, we probably don't want to. And even if we do, modern software will keep a log for us of what we have done (a fact that seems to escape many users of the software). The reason I am not a fan of interactive work has to do with the nature of the analysis we do. The model we ultimately settle on was usually arrived at after several estimates testing many models. And this is appropriate, we ought to all be testing the robustness of our results to changes in model specification, changes in variable measurement, etc.. By writing command files to perform estimates, and keeping each version, one has a record of all of this.[2]

You generally want a large set of comments at the beginning of each file indicating what the file is intended to do. And each file should not do too much. Comments on the top of a file should list the following:

1. State the date it was written, and by whom.
2. Include a description of what the file does.
3. Note what file it was immediately derived from.
4. Note all changes from its predecessors where appropriate.
5. Indicate any datasets the file utilizes as input. If the file utilizes a data-set; there should be a comment indicating the source of the data-set. This could be either another file you have that produced the data; or it could be a description of a public-use data set. If it is a public-use dataset, include the date and version number.
6. Indicate any output files or dataset created.

For example, the first nine lines of a command file might be:

```
\*    File-Name:    mnp152.g
      Date:         Feb 2, 1994
      Author:       JN
      Purpose:      This file does multinomial
                    probits on our basic model.
      Data Used:    nes921.dat (created by
                    mkasc1c.cmd)
      Output File:  mnp152.out
      Data Output:  None
      Machine:      billandal (IBM/RS6000)
*\
```

You could keep a template with the fields above left blank, and read in the template to start each new command file. You should treat the comments at the top of a file the way you would the notes on a table; they should allow the file to stand alone and be interpreted by anyone opening the file without access to other files.

Know the Goal of the Code

It makes no sense to start coding variables if you don't know what the point is. You end up with a bunch of variables that are all being recoded later on and confusing you to no end. Before you start manipulating the data, figure out what you will be testing, and how the variables need to be set up. An

[2] One can go beyond individual command files and create batch-files to run the whole sequence of command files necessary to read the data, process it, and do analysis. There is a lot to be said for this technique; but it is a bit beyond the scope of this article.

example: say you want to test the claim that people who voted in 1992, but did not vote in 1988 were more likely to support Perot than Bush in 1992. How do you code the independent variable indicated here? Well, you really want a 'new voter' variable, since your substantive hypothesis is stated most directly in terms of 'new voters.' Thus the variable should be coded so that:

vote in 1992, not voted in 1988 = 1
voted in 1992, voted in 1988 = 0

Rule: Code each variable so that it corresponds as closely as possible to a verbal description of the substantive hypothesis the variable will be used to test.

Fixing Mistakes

If you find errors, the errors should be corrected where they first appear in your code - not patched over later. This may mean re-running many files. But this is preferable to the alternative. If you patch the error further 'downstream' in the code then you will need to remember to repeat the patch should you make any change in the early part of the data-preparation (i.e., you decide you need to pull additional variables from a data-set, etc). If the patch is downstream you are also likely to get confused as to which of your analysis runs are based on legitimate (patched) data, and which are based on incorrect data.

Rule: Errors in code should be corrected where they occur and the code re-run.

Data-Manipulation Versus Data-Analysis

Most data goes through many manipulations from 'raw-form' to the form we use in our analyses. It makes sense to isolate as much of this as possible in a separate file. For instance, say you are using the 1992 NES presidential election file. You have 100 variables in mind that you *might* use in your analysis. It makes sense to have a program that will pull those 100 variables from the NES dataset, give them the names you want, do any basic recoding that you know you will maintain for all of your analysis, and create a 'system file' of these 100 named and recoded variable that can be read by your statistics package. There are at least two reasons for this. First, it saves a lot of time. You are going to estimate at least 50 models before settling one one. Do you really want to read the whole NES dataset off the disk 50 times when you could read a file 1/20th the size instead? Second, why do all that re-coding and naming 50 times? You might accidentally alter the recodes in one of your files.

Rule: Separate tasks related to data-manipulation vs data-analysis into separate files.

Modularity

Separating data-manipulation and data-analysis is an example of modularity. Modularity refers to the concept that tasks should be split up. If two tasks can be performed sequentially, rather than two-at-a-time, then perform them sequentially. The logic for this is simple. Lots of things can go wrong. You want to be able to isolate what went wrong. You also want to be able to isolate what went right. After what specific change did things improve? Also, this makes for much more readable code.

Thus if you will be engaging in producing some tables before multivariate analysis, you might have a series of programs: descripl.cmd, descript2.cmd, ..., descrip9.cmd. Following this, you might produce: regl.cmd, reg2.cmd,..., reg99.cmd. You need not constrain yourself to one regression per file. But the regressions in each file should constitute a coherent set. For instance, one file might contain your three most likely models of vote-choice - each disaggregated by sex. This does tend to lead to proliferation of files. One can start with regl.cmd and finish with reg243.cmd. But disk-space is cheap these days; and the files can easily be compressed and stored on floppies if things are getting tight.

Rule: Each program should perform only one task.

KISS

Keep it simple and don't get too clever. You may think of a very clever way to code something this week. Unfortunately you may not be as clever next week and you might not be able to figure out what you did. Also, the next person to read your code might not be so clever. Finally, you might not be as clever as you think - the clever way you think of to do 3 steps at once might only work for 5 out of 6 possible cases you are implementing it on. Or it might create nonsense values out of what should be missing data. Why take the chance? Computers are very fast. Any gains you make thru efficiency will be dwarfed by the confusion caused later on in trying to figure out what exactly your code is doing so efficiently.

Rule: Do not try to be as clever as possible when coding. Try to write code that is as simple as possible.

Variable Names

There is basically no place for variables named **X1** other than in simulated data. Our data is real, it should have names that impart as much meaning as we can. Unfortunately many statistical packages still limit us to 8-character names (and for the sake of portability we are forced to stick with 8 character names even in packages that don't impose the limit). However, your keyboard has 84 keys, and the alphabet has 52 letters: 26 lower-case and 26 upper-case. Indulge yourself and make liberal use of them. There are also several additional useful characters -- such as the underscore and the digits 0-9 -- at your disposal. It is a general convention in programming to use UPPER CASE characters to indicate constants, and lower case characters to indicate variables. This might not be as useful in statistical programming. You might adopt the convention that capitals refer to computed quantities (such as **PROBCHC1**: the estimated probability of choosing choice 1). And if you are trying to have your code closely follow the notation of a particular econometrics article, you might use a capital U for utility, or a capital V for the systemic component of utility. Obviously in such a case comments would be in order! Some people like variable names such as **NatlEcR** because the use of capitals allows for clearly indicating where one word stops and another starts. **NatlEcR** makes it easier to think of 'National Economic - Retrospective' than **natlecr** might. You will need to make some tradeoffs in the conventions you choose. The important thing is to adopt a convention on the use of capitals and stick with it.

> **Rule:** Use a consistent style regarding lower and upper case letters.

> **Rule:** Use variable names that have substantive meaning.

When possible a variable name should reveal subject and direction. The simplest case is probably a dummy variable for a respondent's gender; imagine it is coded so that 0 = men, 1 = women. We could either call the variable 'SEX', or 'WOMEN.' It is pretty clear that 'WOMEN' is the better name since it indicates the direction of the variable. When we see our coefficients in the output we won't have to guess whether we coded men=1 or women=1.

> **Rule:** Use variable names that indicate direction where possible.

Similarly, value labels are very useful things for packages that permit them. The examples of computer syntax I use in this article are written in SST (Dubin/Rivers 1992), but they can be easily translated into SAS, SPSS, or most statistical packages. Here is a simple example. The variable **natlecr** indicates the respondent's retrospective view of performance of the national economy. Notice that the variable name can only indicate so much information in 8 characters. But the label of it and the values tell us what we need. And the fact that the label tells us where to look the variable up in the code-book is further protection.

```
label   var[natlecr]     lab[v3531:national
                              economy - retro]
                              val[1 gotbet 3 same \
                              5 gotworse]
```

Some people using NES data -- or any data produced by someone else accompanied by a codebook -- follow the convention of naming the variable by its codebook number (i.e., V3531), and using labels for substantive meaning. I think this is a poor practice. Consider which of the following statements is easier to read:

```
logit      dep[preschc]    ind[one educ
                           women pid]
```

or:

```
logit      dep[V5609]      ind[one V3908
                           V4201 V3634]
```

The codebook name for the variable should *definitely* be retained; but it can be retained in the label statement. Without the codebook name one would not know which of the several party-identification variables the variable **pid** refers to.

Writing Cleanly

Anything that reduces ambiguity is good. Parentheses do so and so parentheses are good. A reader should not need to remember the precedence of operators. But in most cases parentheses are more valuable as visual cues to grouping expressions than to actual precedence of operators. Almost as useful as parentheses is white-space. Proper spacing in your code can make it much easier to read. This means both vertical space between sections of a program that do different tasks, and indenting to make things easier to follow.

Not knowing the syntax of SST, it might be completely opaque that the following code saves all observations of the variables **preschc ... deficit** for which all of the variables contain a valid response. However, the spacing can be a big help towards figuring this out.

```
rem    ************************************
rem    ************************************

recode  var[preschc educ east south west \
            women respfinp natlec \
```

```
          resplib bclibdis \
          gblibdis rplibdis \
          dem rep respgjob resphlth \
          respblk respab \
          age1829 age3044 age4559 \
          newvoter termlim \
          deficit] map[md=-9]


write    var[preschc educ east south west \
          women respfinp natlec bclibdis \
          gblibdis rplibdis             \
          dem rep respgjob resphlth \
          respblk respab  \
          termlim age1829 age3044 age4559 \
          newvoter deficit]            \
          file[nes9212r.asc]   \
        if[(preschc!=-9)&(educ!=-9)& \
          (east!=-9)&(south!=-9)& \
          (west!=-9)&  \
          (women!=-9)&(respfinp!=-9) \
          &(natlec!=-9)&(bclibdis!=-9)& \
          (gblibdis!=-9)&(rplibdis!=-9)& \
          (dem!=-9)&(rep!=-9)& \
          (respgjob!=-9)&(resphlth!=-9)&  \
          (respblk!=-9)&(respab!=-9)& \
          (termlim!=-9)& \
          (age1829!=-9)&(age3044!=-9)& \
          (age4559!=-9)&(newvoter!=-9)& \
          (deficit!=-9)]

rem      *************************************
rem      *************************************
```

> **Rule:** Use appropriate white-space in your programs, and do so in a consistent fashion to make them easy to read.

Kernighan and Plauger (1972) suggest the telephone test, perhaps a bit anachronistic since you will more likely email the code than read it to someone, but useful nonetheless. Read your code to someone over the phone. If they can't understand it, try writing the code again.

Comments

There is probably *nothing* more important than having adequate comments in your code. Basically one should have comments before any distinct task that the code is about to perform. Beyond this, one can have a comment to describe what a single line does. The basic rule of thumb is this: is the line of code absolutely, positively self-explanatory *to someone other than yourself* without the comment? If there is any ambiguity, go ahead and put the comment in.

Notice that in the second sample of code below, the section titled 'Creating Liberal-Distance Variables' does exactly that, it creates a set of variables measuring the ideological distance between respondents and candidates. This would not be so clear from the 8 character variable names. Also notice that the code is a distinct part of the total program and it is easy to see where it begins and ends.

Remember though, the comments should *add* to the clarity of the code. Don't put a comment before each line repeating the content of the line. Put comments in before specific blocks of code. Only add a comment for a line where the individual line might not be clear. And remember, if the individual line is not clear without a comment - maybe you should rewrite it.

> **Rule:** Include comments before each block of code describing the purpose of the code.

> **Rule:** Include comments for any line of code if the meaning of the line will not be unambiguous to someone other than yourself.

> **Rule:** Rewrite any code that is not clear.

Following is a case where a single comment lets us know what is going on:

```
rem     *************************************
rem      Create party-id dummy variables
rem     *************************************

rem      Missing values are handled correctly
         here by SST.
rem      In other statistics packages these
         three variables might
rem      have to be initialized as missing first.

set      dem = (pid < 3)
set      ind = (pid == 3)
set      rep = (pid > 3)

rem     *************************************
rem     *************************************
```

Now here is a case where a longer comment is essential:

```
rem     *************************************
rem      Create ideological-distance variables
rem     *************************************

rem      Ideological distance is computed as
         the distance between the
rem      respondent, and the mean placement
         of the candidate by all
rem      respondents used in our multivariate
         analysis who could place
```

```
rem     the candidate.

rem     The mean-values used here are
        generated by making a first pass
rem     at the data set for the subsample
        we will use here. See the code
rem     at the end of this file for that pass.

rem     This hard-wiring of numbers is
        poor style; but it is very
rem     difficult to automate this
        given the way in which SST handles
rem     missing data.

set     gblibdis = (resplib - 5.32)^2
set     bclibdis = (resplib - 2.98)^2
set     rplibdis = (resplib - 4.49)^2

rem   *************************************
rem   *************************************
```

Most programmers think that well-written code should be self-documenting. This is partly true. But no matter how well-written your code is some comments can make it much clearer.

Recodes and Creating New Variables

Probably the most important thing to keep track of both when recoding variables and creating new variables is missing data. There is no general rule that can specify exactly how to do this, because treatment of missing data can vary across statistics packages. Thus the best rule is:

> **Rule:** Verify that missing data is handled correctly on any recode or creation of a new variable.

In some statistics packages you may be best served by initializing all new variables as missing data, and allowing them to become legitimate values only when they are assigned a legitimate value. The best advice is to recode and create new variables defensively.

> **Rule:** After creating each new variable or recoding any variable, produce frequencies or descriptive statistics of the new variable and examine them to be sure that you achieved what you intended.

Generally it is poor style to 'hard-wire' values into your code. Any specific values are likely to change when some related piece of code somewhere else is altered or when the data-set changes. The example of ideological-distance variables in the Comment section above is an example where values *are* hard-wired into the code. This is a case where a choice had to be made. The values 5.32, 2.98, and 4.49 represent means of candidate placement by a selected set of respondents. Computing this with the **means** command over the appropriate set of respondents, and doing the assignment was complicated enough that rather than write code that could not be simple - it was decided to hard-wire the values. If the following code produced the desired results (it does not), it would be far preferable:

```
rem     *************************************
rem      Create ideological-distance variables
rem     *************************************


set      gblibdis = (resplib - mean(bushlib))^2
set      bclibdis = (resplib - mean(clinlib))^2
set      rplibdis = (resplib - mean(perolib))^2

rem     *************************************
rem     *************************************
```

> **Rule:** When possible, automate things and avoid placing hard-wired values (those computed 'by-hand') in code.

Finally, after you have done recodes and created new variables it is a good idea to list all of the variables. This way you can confirm that you and your statistics package agree on what data is available, and how many observations are available for each variable. In SST this would be done with a list command, in SAS, PROC CONTENTS will produce a clean list of variables. Most statistics packages offer similar commands.

Procedures or Macros

Most political scientists don't ever have to write a 'macro' or 'procedure', but maybe that's why they do so little secondary analysis once they generate some estimates. The purpose of a well-defined procedure is to automate a particular sequence of steps. They are very useful both in making your code more readable, and in allowing you to perform the same operation multiple times on different values or on different variables. The use of Procedures and Macros is really a topic for a separate article; but political scientists should realize that they are available in most statistics packages.

Summing Up

The rules presented here are merely *one way* of accomplishing the goal you should have in mind. That goal is to write clear code that will function reliably and that can be read and understood by you and others and can serve as a road-map for replicating and extending your research.

Most people are in a huge hurry when they write their code. They are either excited about getting the results and want them as fast as possible, or they figure the code will be run once and then thrown out. **If your program is not worth documenting, it probably isn't worth running.** The time you save by writing clean code and commenting it carefully may be your own.

Rules

1. Maintain a labbook from the beginning of a project to the end.
2. Code each variable so that it corresponds as closely as possible to a verbal description of the substantive hypothesis the variable will be used to test.
3. Errors in code should be corrected where they occur and the code re-run.
4. Separate tasks related to data-manipulation vs data-analysis into separate files.
5. Each program should perform only one task.
6. Do not try to be as clever as possible when coding. Try to write code that is as simple as possible.
7. Each section of a program should perform only one task.
8. Use a consistent style regarding lower and upper case letters.
9. Use variable names that have substantive meaning.
10. Use variable names that indicate direction where possible.
11. Use appropriate white-space in your programs, and do so in a consistent fashion to make them easy to read.
12. Include comments before each block of code describing the purpose of the code.
13. Include comments for any line of code if the meaning of the line will not be unambiguous to someone other than yourself.
14. Rewrite any code that is not clear.
15. Verify that missing data is handled correctly on any recode or creation of a new variable.
16. After creating each new variable or recoding any variable, produce frequencies or descriptive statistics of the new variable and examine them to be sure that you achieved what you intended.
17. When possible, automate things and avoid placing hard-wired values (those computed 'by-hand') in code.

Symposium on *Designing Social Inquiry*, Part 1

Larry M. Bartels
Princeton University

King, Keohane, and Verba's *Designing Social Inquiry: Scientific Inference in Qualitative Research* is an important addition to the literature on research methodology in political science and throughout the social sciences.[1] It represents a systematic effort by three of the most eminent figures in our discipline to codify the basic precepts of quantitative inference and apply them with uncommon consistency and self-consciousness to the seemingly distinct style of qualitative research that has produced most of the science in most of the social sciences over most of their history. The book seems to me to be remarkably interesting and useful both for its successes, which are considerable, and for its failures, which are also, in my view, considerable.

Here I shall touch only briefly upon one obvious and very important contribution of the book, and upon one respect in which the authors' argument seems to me to be misguided. The rest of my discussion will be devoted to identifying some of the authors' more notable unfulfilled promises — not because they are somehow characteristic of the book as a whole, but because they are also among the more important unfulfilled promises of our entire discipline. If King, Keohane, and Verba's work stimulates progress on some of these fronts, as I hope and believe it will, their book will turn out to represent a very significant contribution to qualitative methodology.

Anyone who thinks about social research primarily in terms of statistical inference, as I do, has probably thought — and perhaps even said out loud — that the world would be a happier place if only qualitative researchers would learn and respect the basic rudiments of statistical reasoning. King, Keohane, and Verba present those rudiments clearly, engagingly, and with a minimum of technical apparatus. As a result, their work will certainly help to shine the light of basic methodological knowledge into many rather dark corners of the social sciences. For that we owe them profound thanks.

At another level King, Keohane, and Verba's argument seems to me to be misguided, although in a way that seems unlikely to have significant practical consequences. While it is hard to doubt that "all qualitative and quantitative researchers would benefit by more explicit attention to this logic [the logic "explicated and formalized clearly in discussions of quantitative research methods"] in the course of

[1]**Editor's Note:** These comments were originally presented at the 1994 APSA Annual Meetings. The citation for the book is King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton: Princeton University Press.

designing research" (3), it simply does not seem to me to follow that "[a]ll good research can be understood — indeed, is best understood — to derive from the same underlying logic of inference" (4). Even if we set aside theorizing of every sort, from Arrow's (1951) theorem on the incoherence of liberal preference aggregation to Collier's (1994) conceptual analysis of scores of distinct types and subtypes of "democracy," it seems pointless to attempt to force "all good [empirical] research" into the procrustean bed of "scientific inference" set forth by King, Keohane, and Verba. Would it be fruitful — or even feasible — to recast such diverse works as Michels' *Political Parties* (1915), Polanyi's *The Great Transformation* (1944), Lane's *Political Ideology* (1962), Thompson's *The Making of the English Working Class* (1963), and Fenno's *Home Style* (1978) in the concepts and language of statistical inference? Or are these not examples of "good research"?

King, Keohane, and Verba attempt to skirt the limitations of their focus by conceding (43) that "analysts should simplify their descriptions only after they attain an understanding of the richness of history and culture. . . . [R]ich, unstructured knowledge of the historical and cultural context of the phenomena with which they want to deal in a simplified and scientific way is usually a requisite for avoiding simplifications that are simply wrong." But since King, Keohane, and Verba provide no scientific criteria for recognizing "understanding" and "unstructured knowledge" when we have it, the system of inference they offer is either too narrow or radically incomplete. Perhaps it doesn't really matter whether we speak of the process of "attain[ing] an understanding" as a poorly understood but indispensable requisite for doing science or as a poorly understood but indispensable part of the scientific process itself. I prefer the latter formulation, but King, Keohane, and Verba's apparent insistence upon the former will not keep anyone from relying upon — or aspiring to produce — "understanding" and "unstructured knowledge."

Third, and most important, I am struck by a variety of crucial lacunae in King, Keohane, and Verba's codification of good inferential practice. I emphasize these limitations because they seem to me to suggest (though apparently unintentionally) an excellent agenda for the future development of qualitative and quantitative methodology. As is often the case in scientific work, the silences and failures of the best practitioners may point the way toward a discipline's subsequent successes. Here I shall provide four examples drawn from King, Keohane, and Verba's discussions of uncertainty, qualitative evidence, measurement error, and multiplying observations.

One of King, Keohane, and Verba's most insistent themes concerns the importance of uncertainty in scientific inference. They proclaim that "inferences without uncertainty estimates are not science as we define it" (9), and implore qualitative researchers to get on the scientific bandwagon by including estimates of uncertainty in their research reports (9 and elsewhere). But how, exactly, should well-meaning qualitative researchers implement that advice? Should they simply attempt to report their own subjective uncertainty about their conclusions? How should they attempt to reason from uncertainty about the various separate aspects of their research to uncertainty about the end results of that research, if not by the standard quantitative calculus of probability? What sorts of checks on subjective reports of uncertainty about qualitative inferences might be feasible, when even the systematic policing mechanism enshrined in the statistical approach to quantitative inference is routinely abused to the point of absurdity (Leamer 1978; Leamer 1983; Freedman 1983)? Since King, Keohane, and Verba offer so little in the way of concrete guidance, their emphasis on uncertainty can do little more than sensitize researchers to the general limitations of inference in the qualitative mode without providing the tools that would be required to overcome those limitations. As far as I know, such tools do not at present exist; but their development should be high on the research agenda of qualitative methodologists.

King, Keohane, and Verba's discussion of the respective roles and merits of quantitative and qualitative evidence is equally sketchy. While they rightly laud Lisa Martin's (1992) *Coercive Cooperation* and Robert Putnam's (1993) *Making Democracy Work* for combining quantitative and qualitative evidence in especially fruitful ways (5), they provide no clear account of *how*, exactly, Martin's or Putnam's juxtaposition of quantitative and qualitative evidence bolsters the force of their conclusions. Martin's work is rushed precipitously off the stage (as most of King, Keohane, and Verba's concrete examples are), while Putnam's work only reappears — other than in an unrelated discussion of using alternative quantitative indicators of a single underlying theoretical concept (223–224) — in a discussion of qualitative immersion as a source of *hypotheses* rather than *evidence*, which leads to the rather patronizing conclusion that "[a]ny definition of science that does not include room for ideas regarding the generation of hypotheses is as foolish as an interpretive account that does not care about discovering truth" (38).

There is more going on here than a simple-minded distinction between (qualitative) hypothesis generation and (quantitative) hypothesis testing, or than a simple-minded faith that two kinds of evidence are better than one. Qualitative evidence does more than suggest hypotheses, and analyses combining quantitative and qualitative evidence can and sometimes do amount to more than the sum of their parts. King, Keohane, and Verba do little to illuminate those facts. But the larger and more important point is that nobody else does very well, either. Just as the "persuasive force" of such classic works of social science as V. O. Key's (1949) *Southern Politics in State and Nation*, *The American Soldier* (Stouffer et al. 1949), and *Voting* (Berelson et al. 1954) "is not easily explained in conventional statistical theory even today" (Achen 1982, 12), neither is the persuasive force of these and other compelling works convincingly accounted for by partisans of interpretive, ethnographic, historical, or any other brand of qualitative inquiry.

In both of these instances, the limitations of King, Keohane, and Verba's analysis faithfully reflect the limitations of the existing methodological literature on qualitative inference. Some other lacunae in King, Keohane, and Verba's account are attributable to the limitations of the theory of quantitative inference they offer as a model for qualitative research. As a quantitative methodologist — and the coauthor of a rather optimistic survey of the recent literature in quantitative political methodology (Bartels and Brady 1993) — I am chagrined to notice how wobbly and incomplete are some of the inferential foundations "explicated and formalized clearly in discussions of quantitative research methods" (3).

Again, two examples will suffice to illustrate the point. The first is King, Keohane, and Verba's treatment of measurement error, which — like much of the elementary textbook wisdom on that subject — is both incomplete and unrealistically optimistic. They assert that unsystematic (random) measurement error in explanatory variables "unfailingly biases inferences in predictable ways [*sic*]. Understanding the nature of these biases will help ameliorate or possibly avoid them" (155). Later, they assert more specifically that the resulting bias "takes a particular form: it results in the estimation of a weaker causal relationship than is the case" (158). At the end of their discussion they acknowledge that their analysis is based upon a model with a single explanatory variable, but assert that it "applies just the same if a researcher has many explanatory variables, but only one with substantial random measurement error" or if researchers "study the effect of each variable sequentially rather than simultaneously," while allowing that "if one has multiple explanatory variables and is simultaneously analyzing their effects, and if each has different kinds of measurement error, we can only ascertain the kinds of biases likely to arise by extending the formal analysis" (166).

King, Keohane, and Verba's assertion about the case of several explanatory variables where only one is measured with substantial error is quite misleading, in that it fails to note that the bias in the parameter estimate associated with the one variable measured with substantial error will be propagated in complicated ways to all of the other parameter estimates in the analysis, biasing them upward or downward depending upon the pattern of correlations among the various explanatory variables. Their assertion about sequential rather than simultaneous analysis of several explanatory variables is also misleading, at least in the sense that the resulting omitted variable bias may mitigate, exacerbate, or reverse the bias attributable to measurement error. And the promise of "ascertain[ing] the kinds of biases likely to arise" in more complicated situations "by extending the formal analysis" can in general be redeemed only if we have a good deal of prior information about the nature and magnitudes of the various errors — information virtually impossible to come by in all but the most well-understood and data-rich research settings (Achen 1983; Cowden and Hartley 1993). Thus, while it seems useful to have alerted qualitative researchers to the fact that measurement error in explanatory variables may lead to serious biases in parameter estimates, it seems disingenuous to suggest that statistical theory offers reliable ways to "ameliorate or possibly avoid" those biases in real qualitative research.

King, Keohane, and Verba's chapter on "Increasing the Number of Observations" seems equally disingenuous in asserting that "almost any qualitative research design can be reformulated into one with many observations, and that this can often be done without additional costly data collection if the researcher appropriately conceptualizes the observable implications that have already been gathered" (208). While they are right to emphasize the importance of "maximizing leverage" by using the available data to test many implications of a given theory (or even better, of several competing theories), their discussion obscures the fact that having many *implications* is not the same thing as having many *observations*. In order for our inferences to be valid, each of our many implications must itself be verified using a research design that avoids the pitfall of "indeterminacy" inherent in having more explanatory variables than relevant observations.

What, then, *is* a "relevant observation"? King, Keohane, and Verba themselves provide the answer in their earlier, clear and careful discussion of "unit homogeneity." Relevant observations are those for which "all units with the same value of the explanatory variables have the same expected value of the dependent variable" (91). But the more we succeed in identifying diverse empirical implications of our theories, the less likely it will be that those diverse implications can simply be accumulated as homogeneous observations in a single statistical model. Having a richly detailed case study touching upon many implications of the same theory or theories is no substitute for "seek[ing] homogeneous units across time or across space" (93), as King, Keohane, and Verba themselves point out in their subsequent discussion of "process tracing" (226–228).

King, Keohane, and Verba allow that "[a]ttaining unit homogeneity is often impossible," but go on to assert in the next sentence that "understanding the degree of heterogeneity in our units of analysis will help us to estimate the degree of uncertainty or likely biases to be attributed to our inferences" (93–94). How is that? Again, they do not explain. But again, the more important point is that nobody else does, either — a point I am compelled to acknowledge despite my own recent efforts in that direction (Bartels 1994). If we accept King, Keohane, and Verba's assertion that the "generally untestable" assumption of unit homogeneity (or the related assumption of "constant causal effects") "lies at the base of all scientific research" (93), that is a loud and embarrassing silence.

In the end, King, Keohane, and Verba's optimistic-sounding unification of quantitative and qualitative research seems to me to promise a good deal more than it delivers, and a good deal more than it could possibly deliver given the current state of political methodology in both its qualitative and quantitative modes. But perhaps that is the genius of the book. By presenting a bold and beguiling vision of a seamless, scientific methodology of social inquiry, King, Keohane, and Verba may successfully challenge all of us to make some serious progress toward implementing that vision.

References

Achen, Christopher H. 1982. *Interpreting and Using Regression.* Beverly Hills: Sage Publications.

Achen, Christopher H. 1983. "Toward Theories of Data: The State of Political Methodology." In *Political Science: The State of the Discipline,* ed. Ada W. Finifter. Washington, DC: American Political Science Association.

Arrow, Kenneth J. 1951. *Social Choice and Individual Values.* New Haven: Yale University Press.

Bartels, Larry M. 1994. Pooling Disparate Observations. Paper presented at the Annual Meeting of the American Political Science Association, New York.

Bartels, Larry M., and Henry E. Brady. 1993. "The State of Quantitative Political Methodology." In *Political Science: The State of the Discipline II,* ed. Ada W. Finifter. Washington, DC: American Political Science Association.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

Collier, David. 1994. Classical and Radial Categories in Comparative Analysis: Extensions and Implications. Paper presented at the Annual Meeting of the American Political Science Association, New York.

Cowden, Jonathan A., and Thomas Hartley. 1993. "Complex Measures and Sociotropic Voting." In *Political Analysis,* Volume 4, ed. John R. Freeman. Ann Arbor: University of Michigan Press.

Fenno, Richard F. 1978. *Home Style: House Members in Their Districts.* Boston: Little, Brown.

Freedman, David A. 1983. "A Note on Screening Regression Equations." *American Statistician* 37: 152-155.

Key, V. O., Jr. 1949. *Southern Politics in State and Nation.* Knoxville: University of Tennessee Press, 1984.

Lane, Robert E. 1962. *Political Ideology: Why the American Common Man Believes What He Does.* New York: Free Press.

Leamer, Edward E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data.* New York: John Wiley & Sons.

Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." *American Economic Review* 73: 31-43.

Martin, Lisa L. 1992. *Coercive Cooperation.* Princeton: Princeton University Press.

Michels, Robert. 1915. *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy.* Eden and Cedar Paul, trans. New York: Free Press, 1962.

Polanyi, Karl. 1944. *The Great Transformation: The Political and Economic Origins of Our Time.* Boston: Beacon Press, 1957.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton: Princeton University Press.

Stouffer, Samuel A., et al. 1949. *The American Soldier.* Princeton: Princeton University Press.

Thompson, E. P. 1963. *The Making of the English Working Class.* New York: Vintage Books, 1966.

Symposium on *Designing Social Inquiry*, Part 2

Henry E. Brady
University of California, Berkeley

Doing Good and Doing Better

What kind of contribution is *Designing Social Inquiry* by Gary King, Bob Keohane, and Sidney Verba?

Theological seminaries distinguish between theology, or the systematic study of religious beliefs, and homiletics, the art of preaching the gospel convincingly. Theologians ask hard questions, develop new systems of theology, and often espouse opinions that would shock and horrify the practicing and devout members of the religion's congregations. Homiletics is about homilies—practical, down to earth, simple, and above all, reliable, interpretations of the faith. Religions understand, as the social sciences may not, that the goal is to save souls and not simply to increase our knowledge or understanding of the world. For this reason, both theology and homiletics have pride of place in seminaries.

The social sciences have a great deal of theology, but very little homiletics. Perhaps this is why we have saved so few souls. And it may also be why we do such a bad job training students. A little homiletics might go a long way towards improving our discipline.

Designing Social Inquiry (DSI) is a homily, not theology. There is art in a good homily. Like all good homiletic literature, *DSI* puts aside doubt and complexity. After all, who would want to burden the average graduate student with the tedious complexity of St. Thomas Aquinas in *Summa Theologica* or Paul Tillich in *Systematic Theology*? And who would recommend the self doubt of St. Augustine's *Confessions* or Kierkegaard's *Fear and Trembling, or The Sickness Unto Death*? Better to give them Norman Vincent Peale's *The Power of Positive Thinking.*

Designing Social Inquiry, however, is not just about positive thinking. It is closer to Moses Maimonides *Guide for the Perplexed* or Luther's *A Catechism for the People, Pastor and Preacher*. It has a powerful message about the need for reform, self-sacrifice, and discipline on the part of all political scientists—especially qualitative researchers.[1] It puts forth a simple, straightforward faith. And it tries very hard not to treat qualitative researchers as souls unworthy of salvation. To the contrary, it envisions a unified social science in which there are "Two styles of research, One Logic of Inference."[2] To practice this one logic of inference, *DSI* presents a simple unified series of steps, a faith to live by, based upon insights from statistics and econometrics. In Chapter 3, for example, we are told to:

- Construct falsifiable theories
- Build theories that are internally consistent
- Select dependent variables carefully
- Maximize concreteness
- State theories in as encompassing a way as possible.

In homiletic literature, exhortations, such as these, should be simple, and they need not always be completely consistent (witness the last two rules listed above). A good sermon should have clear points; it should avoid doubt; it should provide plenty of examples. The goal should be to convert the heathen qualitative researcher to the true faith.

This book — TO ITS CREDIT — does these things. It is an extraordinarily good piece of homiletic literature and it should be used in the classroom. It is very nicely written. It is generally lucid and well-organized. No one can fail to hear its message.

We should all probably hear the message that is preached. I, for one, have great sympathy with this enterprise, having spent far too many hours listening to talks on comparative politics in which dependent variables or independent variables (or both) did not vary, in which selection bias seemed insurmountable, in which explanations seemed more like good stories than hard won insights from ruling out alternative possibilities. I too, in my introductory statistics classes, have tried to point out to comparativists that they could do so much better if they avoided omitted variable bias, stopped selecting on the dependent variable, and so forth. I have used some of the same diagrams displayed in the text (e.g., Figures 4.1, 5.1, and 5.2) to make didactic points about good research.

Why, then, do I find myself worried about what this book tries to do? Perhaps I am worried because despite the book's desire for a unified approach to social science, there may be something wrong with quantitative researchers,[3] who luxuriate in large numbers of observations and even the possibility under some circumstances of doing experiments, trying to impose a code of conduct, a morality, taken from their own experiences. Certainly the authors, three of the most distinguished and intelligent political scientists in our discipline, mean well, think well, and write well. But I worry that, in the end, they are a little like the Reverend Ike who when asked how he reconciled living in luxury while he preached to the poor responded that he believed that the best thing you could do for the poor was not to be one of them. The book ends, in fact, with a chapter on "Increasing the Number of Observations."[4] Is this the best thing we can do for qualitative researchers: to recommend that they not be "small-n" researchers?[5]

[1] *Designing Social Inquiry*, subtitled "Scientific Inference in Qualitative Research," begins by discussing the relationship between quantitative and qualitative research, but another dichotomy also runs through the book. Quite often the authors are more concerned with juxtaposing "small n" versus "large n" research than with the qualitative-quantitative distinction. These are not the same things. Small n research is often qualitative, but it need not be, and large n research can be qualitative. Roughly speaking, the qualitative-quantitative distinction revolves around issues of concept formation and measurement whereas the small n versus large n distinction brings up problems of defining the relevant populations, sampling from them, and dealing with statistical variability. I argue later in this review that these statistical issues are dealt with much more clearly in the book than are those regarding concept formation and measurement.

[2] This phrase resonates especially well with someone like myself who was brought up as a Catholic where the faithful must deal with the mystery of three manifestations of God (in the Father, Son, and Holy Spirit) in a monotheistic religion. By childhood training, I am quite receptive to a message of mono-methodism, even in those circumstances where it requires a leap of faith.

[3] Bob Keohane is not a quantitative researcher, but two of the authors, Gary King and Sidney Verba certainly are, and the book's approach is so rooted in quantitative research that it seems fair to make this assertion.

[4] The chapter means more and does more than just suggest that qualitative researchers get more data, although that is one of the recommendations. I make more comments about this interesting chapter later in the review.

[5] Indeed, qualitative researchers may profit by increasing their observations, and one of the great strengths of this book is that it tries to indicate how the poor in observations can become richer in their understanding. At the same time, the book's unspoken presumption that qualitative researchers are inevitably handicapped

Descending from the Rhetorical Heights

Some readers may think it unfair to judge a book by the vague worries it arouses. We should, it might be argued, judge a book by its logic and its argument. This I now try to do, but before doing so, it may be worth remarking that the value of the *Baltimore Catechism* in which I was drilled as a child should not be measured by its logic and argument. Rather it should be evaluated in terms of how many children it saved from perdition. In the end, I think that is how *DSI* should be judged. Does it work in a classroom? Does it make us better social scientists? In fact, I think it rather unfair to be too scrupulous, as I am about to be, about the details of its arguments.

I have a number of concerns with *DSI*. Here I will focus on two: My belief that *DSI* is handicapped by a "statistician's" view of causality and my desire to see more discussion of measurement problems.

Explanation and Causality

After a useful discussion of descriptive inference or "establishing facts" in Chapter 2, King, Keohane, and Verba (KKV) go on in Chapter 3 to discuss "Causality and Causal Inference." As far as I can tell, they equate explanation with causal thinking.[6] Yet philosophers of science are not so sure that the only kind of explanation involves causality. Take, for example, "classification" explanations such as the observation that iron has certain properties because it appears in a certain column of the periodic table. This does not appear to be a causal explanation.[7] It could be argued that Bohr's atomic theory and its extensions in modern quantum mechanics has provided a causal explanation, but this only amounts to saying that there may be causal explanations as well as classification explanations. Moreover, there was a substantial period of time when the classification explanation was all we had. Should we discard these explanations, even when they are all we have, because they do not appear to be causal? We are not so rich with explanations in the social sciences that we can afford to do this without good reason. Qualitative social scientists, in fact, seem especially fond of typologies and classification systems. Are these worthless? I do not personally have an answer to my question, so perhaps I should not fault KKV for failing to include a discussion of this difficult issue. But it is perplexing and thought provoking.

The approach to causality advanced in *DSI* is based upon an interesting framework developed by the statisticians Donald Rubin (1974; 1978) and Paul Holland (1986). The great strength of this approach, to my mind, is that it emphasizes that a definition of causality requires (a) the careful description of a couterfactual condition (what would have happened if the cause had been absent?) and (b) a comparison of what did happen with what would have happened had the cause been absent. These are two powerful points, and DSI is to be commended for bringing them to the forefront of our discussion. Researchers of all stripes should spend more time describing the couterfactual world that underlies their "becauses". What does it mean, for example, to say that "turnout is lower in that district because it has a high proportion of minorities"? What is the couterfactual world in which turnout would be higher? Is it simply one with a lower proportion of minorities? Would these non-minorities be like minorities in every other respect except race? How could this happen? What would it mean to have it happen?[8] These are not easy questions.

by lack of quantification and small numbers of observations is bothersome. It ignores the possibility that quantitative researchers may sometimes be handicapped by procrustean quantification and a jumble of dissimilar cases.

[6] It is not exactly clear how "explanation" fits into KKV's categories of descriptive and causal inference but one reasonable interpretation is that they consider explanation to be identical with causal inference. In the first three paragraphs of Chapter 2, KKV repeatedly refer to the "dual goals of describing and explaining (34)." They also note that "Description and explanation both depend upon rules of scientific inference. In this chapter we focus on description and descriptive inference (34)." This suggests that Chapter 3, on "Causal Inference" is about explanation. Yet, things cannot be quite so simple because they go on to say that "As should be clear, we disagree with those who denigrate 'mere' description. Even if explanation—connecting causes and effects—is the ultimate goal, description has a central role in all explanation, and it is fundamentally important in and of itself." The first part of the sentence seems to define explanation as "connecting causes and effects", but the later part seems to say that description is also a form of explanation. In the sentence after this one, KKV retain the duality of description and explanation and seem to equate explanation with causal inference, but they argue for the primacy of inference over either one: "It is not description versus explanation that distinguishes scientific research from other research; it is whether systematic inference is conducted according to valid procedures. Inference, whether descriptive or causal, qualitative or quantitative, is the ultimate goal of all good social science."

[7] For more discussion of this example and whether there are noncausal explanations, see Chapter 7 of Achinstein (1983). Brody and Grandy (1989) provide an excellent set of readings on these topics. Gary King has suggested (personal communication) that classification is a form of descriptive inference, but this seems to stretch KKV's concept of descriptive inference beyond distinguishing "the systematic component from the non-systematic component of the phenomena we study" (56). It also adds to the confusion noted in the preceding footnote.

[8] I have deliberately chosen an example in which the putative cause is a characteristic that might be thought unchangeable. Holland, for example, argues that is impermissible to call race or gender a cause because "For causal inference, it is critical that each unit be potentially exposable to any one of the causes. As an example, the schooling a student receives can be a cause, in our sense, of the student's performance on a test, whereas the student's race or gender cannot." (Holland, 1986, 946). This point is not much in evidence in DSI, and I think KKV were wise to minimize its importance because it certainly seems possible to imagine a world in which gender or race changes but nothing else.

I have already argued that there might be explanation without causality. I think there might also be causal effects without (much) explanation. Suppose we find, to use KKV's example, that incumbent legislators do better in elections than non-incumbent legislators. Suppose, in fact, we are as certain as we can be about this because we have done an experiment ("random term limits," for example) with a large N to test it out. This finding immediately leads to other questions about what aspects of incumbency create this advantage (see, for example, Cain, Ferejohn, and Fiorina, 1987). These questions amount to a desire to further specify the causal mechanism. KKV are not adverse to specifying causal mechanisms, and they say that "any coherent account of causality needs to specify how the effects are exerted,"[9] but they believe that "our definition of causality is logically prior to the identification of causal mechanisms." (86) This claim of logical priority may or may not be true (I'm not sure it is very important), but what is true, it seems to me, is that a discussion of causality is inevitably tied up with a discussion of explanation, theories, and causal mechanisms, and *DSI* does not pay enough attention to this relationship. There is no discussion of Hempel's (1965) covering laws, of Wesley Salmon's (1984) model of statistical explanation, of Scriven's (1975) "Causation as Explanation," and many other important works on this topic. This is surprising because the philosophical literature, at least, cannot seem to separate the discussion of these issues.[10]

The statistics literature, in fact, is exceptional in defining causality without discussing explanation. Perhaps this is because statisticians want a method of inference that relies only upon the research design and the data and not at all upon the substance of the research. Yet the net result of the Holland-Rubin papers is a definition that seems surprisingly distant from the problems of theory building and explanation as it exists in the sciences. Most importantly, this approach provides no guidance on what constitutes a "good" explanation beyond what constitutes a good causal inference. Yet an analysis of the impact of incumbency may be an excellent causal inference while being a bad explanation.[11]

After defining causality, KKV go on to describe a method for causal inference. In this, as in their definition, they are guided by the work of Holland and Rubin. The major strength and weakness of this approach is its reliance upon the metaphor of the controlled experiment for solving the problem of causal inference. Holland tells us that "because experimentation is such a powerful scientific and statistical tool and one that often introduces clarity into discussions of specific cases of causation, I unabashedly draw on the language and framework of experiments for the model for causal inference. It is not that I believe an experiment is the only proper setting for discussing causality, but I do feel that an experiment is the simplest such setting." (1986, 946). Fair enough. But it is worrisome that Holland finds it "beyond the scope of this article to apply the model for causal inference to nonrandomized studies" (949), and he cites other literature (Rubin, 1978) that essentially concludes that nonrandomized studies are exceptionally difficult to analyze. And it is telling that Rubin's (1978) extension of the basic framework requires modeling "(1) the prior distribution of the potentially observable data, (2) the mechanism that selects experimental units for exposure to treatments and assigns treatments, and (3) the mechanism that chooses values to record for data analysis." (35) This is a lot of modeling, and it only seems possible if we have some strong theories to draw upon.

KKV provide a simplified version[12] of the Holland and Rubin framework, and in the process ignore some of its subtleties. The crucial part of the book is their discussion of "Conditional Independence" on pages 94-96. In the Holland-Rubin set-up there are as many dependent random variables as there are variations in the treatment condition or the explanatory variable(s). In the simplest case with two levels of the treatment, this implies two random variables. One describes the values on the dependent variable Y for the situation where all cases in the population[13] assume that get one level of the treatment (call this Y^I to match KKV's terminology) and the other is for the values on the dependent variable for the situation where all cases in the population cases get the other level of the treatment (call this Y^N and assume for simplicity that it is no treatment at all). In the real world and for any feasible design, at least one of these values must be censored for each case. That is, we cannot give a case some treatment and no treatment

[9] KKV present this sentence on pages 85 and 86.

[10] Brody and Grady (1989), for example, link them in Part II of their reader entitled "Explanation and Causality." Scriven (1975) joins the two concepts in his famous article on "Causation as Explanation," and every philosophical writer of whom I am aware deals with explanation and causation together.

[11] If the incumbency example does not persuade, consider a doctor called upon to explain the incidence of psychedelic experiences in a remote culture. In an experiment, the doctor shows that a treated group eating a plant diet consisting of peyote, hemp, beans, carrots, and other plants has a statistically significant increase in their incidence of psychedelic experiences. Thus, eating plants causes psychedelic experiences. This is clearly an incomplete explanation. I wish KKV had discussed by what method I might improve it. I think a discussion of a "good explanation" that went beyond methods for finding causal impacts would have gone a long way towards solving this problem.

[12] They do add one complexity by making a useful distinction between "realized causal effect" and "random causal effect," but they suppress so much notation and philosophical discussion in their presentation that many of the nuances in the Holland's (1986) presentation are lost and none of the extensions in Rubin (1978) are discussed.

[13] In this exposition I ignore sampling problems by assuming observations are available on all members of the population. If the entire population cannot be observed, then some assumption has to be made about random sampling.

at the same time. But Y^I and Y^N are not the censored variables; they include the unobserved (and unobservable) values as well as the observed ones. A reasonable definition of the causal effect of the treatment is the average of Y^I minus the average of Y^N, but this quantity cannot be calculated because of the unobserved values in these two random variables.

In the Holland-Rubin framework, a sufficient assumption for estimating a causal effect is the assumption of independence between the assignment of treatments and the random variables Y^I and Y^N. This insures, for example, that people who are high on Y^I are not more likely to get a high level of treatment than those who are low on Y^I. Consequently, we can be sure, for a large enough sample, that the difference in the average of the dependent variable for those who did get the treatment —this quantity can be calculated—and the average in the dependent variable for those who did not get the treatment (another calculable quantity) is the size of the causal effect. One way to achieve this kind of independence is to have correctly carried out randomized experiments.

KKV's discussion of this is a bit opaque, and they seem to conflate the independence assumption with conditional independence.[14] Conditional independence is the assumption that the values of Y_I and Y_N conditional on "pre-exposure" or "control" variables are independent of the assignment of treatments. This is implied by independence but it is a much less stringent assumption. It is the assumption that is usually required for the analysis of quasi-experiments (Achen, 1986). The conflation of these two different assumptions creates difficulties in the exposition because whereas we have a method, random assignment to treatment, for attaining independence, we have no comparable method for insuring that the conditional independence assumption holds outside of a randomized design. The best we have is the checklist of "threats to internal validity" developed by Donald Campbell with Julian Stanley and Thomas Cook (Campbell and Stanley, 1966 and Cook and Campbell, 1979).[15] The rest of *DSI* can be considered another approach to developing a checklist of threats to validity.

Unfortunately, KKV do not allow themselves enough pages in this short section to make this very important transition from a discussion of causal inference for experiments to causal inference with "quasi-experiments." I wish they had taken more time to explain the independence assumption in detail and to show how randomized experiments might provide us with an operational procedure that would make this assumption plausible. In doing this, they would no doubt have come to the conclusion presented by Cook and Campbell (and updated and expanded recently by Heckman, 1992) that there are many reasons to worry about the efficacy of randomization when humans are involved. There are numerous ways in which human beings can make the treatment endogenous by changing their behaviors. There are additional problems when drop-outs (and hence censoring of observations) vary by treatment. And there are the difficulties of truly randomizing units when they are people or groups. Once these problems are recognized for randomized designs, it becomes easier to understand how difficult it is to insure conditional independence for non-randomized designs.

This transition section might also benefit from a more careful discussion of how theories provide the fundamental basis for making a claim of conditional independence. This is an extraordinarily important step, and knowing how to do it can help researchers avoid the inferential nihilism that has crept into some statisticians' discussion of causal thinking in the social sciences (e.g., Freedman, 1991). According to this line of thinking, randomized experiments are practically the only reliable way to be confident that the conditions for reasonable inferences are met. Conditional independence is considered a chimera—seldom justifiable and usually accepted by the researcher as a matter of pure faith and nothing more. Indeed, if I accepted a notion of inference as bare of theories and the logic of explanation as that proposed by Rubin and Holland, I might also be sceptical of conditional independence. But I believe that it is possible to use our prior knowledge, our theories, to carry out the three modeling steps laid out by Rubin (and cited above) so that I am more sanguine about the possibilities for cautiously asserting conditional independence.

It might be argued that I brood unnecessarily over technical points. But the section on "Conditional Independence" is the linch-pin of this book. KKV want to show us that concepts from statistics and econometrics will improve our ability to do qualitative research. They argue that the

[14] This accounts for the confusing set of sentences at the beginning of section 3.3.2 where KKV first say that "Conditional independence is the assumption that values are assigned to the explanatory variables independently of the values taken by the dependent variables" and then go on to say "That is, after taking into account the explanatory variables (or controlling for them), the process of assigning values to the explanatory variable is independent of both (or, in general two or more) dependent variables, Y_i^N and Y_i^I." The first quoted sentence must refer to the independence assumption (because conditional independence does not assume that the values assigned to the explanatory or control variables are independent of the values of the dependent variables) whereas the second quoted sentence appears to be about conditional independence.

[15] I was surprised to find that none of Campbell's publications were referenced in *DSI*. Besides the books referenced in the text, Campbell's selected papers on *Methodology and Epistemology for Social Science* (1988) make excellent reading.

essence of good social research is establishing causal effects. This, in turn, requires making an assumption about conditional independence. This assumption, they believe, can be made plausible by avoiding clear-cut violations of it described in the statistics literature. Yet at the crucial transitional moment the argument seems muddy to me. Exactly how can we rule out the violations identified by quantitative researchers? Do quantitative researchers do a good job in this regard? How sure can I be that conditional independence holds after I have followed the instructions in DSI?

DSI goes on to make many useful observations (although, to be honest, I think that Don Campbell and his collaborators have more useful lists of threats to validity and more trenchant comments about the problems of doing quasi-experimental research), but the crucial argument seems to be missing.

Measurement

KKV devote about fifteen pages to measurement. About three pages cover the "nominal, ordinal, interval" distinction found in the classic papers by S.S. Stevens (1946, 1951), and the remaining twelve are about systematic and nonsystematic measurement error. The major results on measurement error are the classic ones dating from at least Tintner (1952) on how error in the dependent variable does not bias regression results whereas error in the independent variable produces bias in regression coefficients—in fact, biases them unambiguously downwards in the bivariate case. These are well-known results, often repeated in one form or another in classic primers on research design such as Kerlinger (1979), but I do not think they get at the heart of what can be learned from the extensive literature on measurement.

Designing Social Inquiry probably gives such short shrift to measurement because KKV believe that causal inference, roughly what Cook and Campbell call "internal validity," is the central problem of doing good social science. I trace this belief to their decision to equate explanation with causal thinking, and to define causal thinking in narrow statistical terms. Through this progression, the problems of theory construction, concept formation, and measurement recede into the distance. Yet it seems to me that concept formation and measurement, or what Cook and Campbell call the problem of "construct validity," are equally important in almost all instances and possibly of paramount importance in qualitative research. Certainly notions such as "civil society," "deterrence," "democracy," "nationalism," "material capacity," "corporatism," "group-think," and "credibility" pose extraordinary conceptual problems just as "heat," "motion," and "matter" did for the ancients. It may be comforting for the qualitative researcher to know that the true effects of these error-laden variables are even larger in magnitude than what we would estimate using a standard regression equation, but most qualitative researchers are struggling with much more basic problems such as figuring out what it means to measure their fundamental concepts. These problems are certainly not solved by telling us to decide whether the concept is nominal, ordinal, or interval and by admonishing us to "use the measure that is most appropriate to our theoretical purposes." (153).

I will not pretend to have the answers to the problems of construct validity in qualitative research, but I think that the debates on these problems would have been advanced by citing some of the more recent literature in this area. Among the notions that come to mind, let me mention three topics that might have been included. Something might have been said about the conceptualizations of measurement developed by Krantz, Luce, Suppes, and Tversky in their magisterial three volume work on *Foundations of Measurement* (1971-1990), the related notions put forth by Georg Rasch in his quirky but very influential work on *Probabilistic Models for Some Intelligence and Attainment Tests* (1960), and the fascinating *Notes on Social Measurement* (1984a) penned by Otis Dudley Duncan who followed-up this broadside on the limitations of social measurement with a brief for using Rasch models in the social sciences (1984b) . These works show that *qualitative* comparisons are the basic building blocks of any approach to measurement, thus bridging the "quantitative-qualitative" divide by showing that the two approaches are intimately related to one another. This discussion would have easily led to a second topic: the dimensionality of concepts, the nature of similarity judgments that often underlie concept formation, and the role of taxonomies and classifications in science. Finally, there might have been a survey of how the LISREL framework (Bollen, 1991), especially when it is combined with the "multi-trait multi-method" of Campbell and Fiske (1959), sheds light on the practical problems of measurement.

Let me discuss each of these literatures. Duncan's observations on Stevens' scale types are probably the best starting place:

> I conclude that the Stevens theory of scale types, pruned of its terribly misleading confusion of classifications and binary variables with N scales, augmented to take more explicit account of the scales used in measuring numerousness and probability, and specified more clearly so that the examples could be properly understood and assessed, has utility in suggesting the appropriate mathematical and numerical treatment of numbers arising from different kinds of measurement. *Still, a theory of scale types is not a theory of measurement.* And I, for one, am doubtful that any amount of study devoted to either of those topics can teach

> you how to measure social phenomena, though it can conceivably be helpful in understanding exactly what is achieved by a proposed method of measurement or measuring instrument. (1984a, 154, Italics added)

Lest anyone miss Duncan's point, his next chapter is entitled "Measurement: The Real Thing." What is "the real thing?"

Krantz, Luce, Suppes, and Tversky (1971-1990) provide the fullest answer to this question, but Duncan provides a more accessible treatment. Measurement, Duncan argues, is not the same as quantification, and it must be guided by theories which emphasize the relationships of one measure to another. Take, for example, that favorite illustration of introductory methods classes, the measurement of temperature. Although the development of thermometry involves a complicated interplay between theory and invention, one of the important milestones was the discovery of the gas law for which temperature is proportional to pressure times volume. Thermometry only began to progress beyond crude ordinal distinctions such as cold, warm, or hot to true interval scales once laws like the gas law made it clear that temperature could be measured by the change in volume of some material under constant pressure.

One of the distinctive features of this way of measuring temperature is that it relies upon a simple multiplicative law which relates temperature to two quantities which can be "extensively" measured. Extensive measurement refers to the use of the standard millimeter, gram, second, or some other quantity which can be duplicated so that a number of them can be added together ("concatenated") and compared with some object or phenomenon whose length, weight, duration, or other feature is unknown. There is no such standard for temperature, but it can still be measured because it is related to two quantities which can be measured extensively. (Volume as the cube of length and pressure as mass times length per time and area squared.)

A fundamental difficulty facing empirical social science is the apparent impossibility of developing extensive measurements of many important theoretical quantities. Consider, for example, the notion of utility which is basic to both economics and public choice theory in political science. Utility cannot be measured extensively, but economists avoid this difficulty through an ingenious ploy: They throw utility out of their empirical models by deriving demand curves from the maximization of utility with respect to a budget constraint that consists of the sum of prices times quantities. This produces a demand curve—an equation in prices and quantities—both of which can be measured extensively. This ploy, unfortunately, does not appear to be readily available to political scientists.

The contribution of Georg Rasch (1960; Andrich, 1988) and of Krantz, Luce, Suppes, and Tversky (1971) in their method of "conjoint measurement" has been to show how measurement can be carried out without an extensive measure that can be duplicated and combined: all that is needed is the ability to make *qualitative* distinctions about the amount of each of several variables which are thought to be multiplicatively related to one another. Rasch's method, designed for scoring achievement tests, has the great virtue that it scores both test-takers and the items on the test simultaneously.

All this fancy talk does not provide us with a straightforward way to measure the basic concepts in qualitative social science, but it provides us with some clues about how we might go about measuring these concepts. First, it suggests that we have two basic strategies for measurement. We can either try to define a concept extensively (as with length, weight, prices, or quantities) or conjointly (as with achievement tests and subjective probability). Thus we can measure democracy extensively by the fraction of the population enfranchised or by the number of parties, or we can measure it conjointly by using ratings from knowledgeable observers. If we use the second method, and qualitative researchers might be inclined to use this second approach, then we might want to think about whether we should scale the raters as well as the countries that are rated. Maybe raters differ in their willingness to call a country a democracy; maybe they even have biases of some sort or another. Second, this discussion suggests that theories must help to guide the measurement process. In their impressive series of papers on bias in electoral systems, Gelman and King (1994) follow just this strategy with a simple framework for thinking about representation. Steven Fish (1995) also does this (more implicitly than explicitly) in his discussion of the development of civil society in Russia. One of his indicators of civil society is the aggregation of interests by groups which he describes as the group's "identification of 'cleavage issues' and the formulation of specific goals and agendas [and] ... the formation of a collective identity, which includes the identification of a membership (53-54)." Although Fish does not provide a mathematical description of his measure, it could be conceptualized as the degree to which participation or membership in a group is highly correlated with some politically relevant characteristic or cleavage. This amounts to defining this component of civil society as the product of group participation and a politically relevant characteristic—a multiplicative relationship of the sort described by measurement theorists as indicative of true measurement. Fish's approach makes sense partly because it has exactly this form. Hence, measurement theory provides a clear-cut check on when we can say that we have the framework for measuring something.[16]

[16]Gary King (personal communication) suggests that these are points for quantitative researchers and not qualitative researchers because they deal with quantitative measures. Putting aside the fact that a discussion of measurement error or Stevens' scale types

This approach leads immediately into the next topic I mentioned above. There is a very rich literature on the "topology" of measurement that indicates what is required for single or multi-dimensional measures; what is required for dimensionality itself; what is required before something is considered the same as something else; and under what conditions objects can be better taxonomized using "trees" or Euclidian space. These methods are now widely used in biology to inform studies of evolution. I suspect that they would be quite useful for the qualitative researcher who wants to trace the evolution of the concept of democracy over time, or the similarities and differences among contemporary democracies.[17] After all, qualitative researchers often spend a great deal of time and effort developing typologies and taxonomies.

Finally, although I often worry about the wholesale use of LISREL in survey research, I think the marriage of factor analysis to simultaneous equation modeling in LISREL has made many researchers more aware of measurement problems. Ken Bollen (1993) presents an exemplary use of the technique for analyzing three sets of ratings by three different people for political liberties and democratic rule in countries around the world. By having two concepts in mind, Bollen is able to search for "discriminant" as well as "convergent" validity as Campbell and Fiske (1959) tell us we should do. Bollen also allows for the possibility that raters may have biases, and he finds, for example, that one rater "tends to favor countries in Central America and South America, Western Industrial nations, and, to a lesser extent, countries in the Oceania region" while providing lower scores for Sub-Saharan Africa, Eastern Europe and Asia regions. One can imagine extending Bollen's work by adding other methods for rating democracy and by examining (as he does in a preliminary way) how the characteristics of the raters affect their ratings. Bollen's work suggests that qualitative researchers might improve their understanding of concepts by considering various definitions of them, by considering concepts closely related to them, and by considering concepts that are different from them. This strategy, for example, is followed by Hannah Pitkin in her classic work on representation (1967). [18]

assumes the same thing (and the entirety of *Designing Social Inquiry* is based upon the premise that quantitative methods provide lessons for qualitative researchers), it is worth noting that qualitative researchers also engage in comparisons that amount to a form of measurement. Qualitative researchers should know that quantitative research relies upon just the kinds of comparative statements that are at the core of qualitative research. In fact, a discussion of this sort would lead to a conclusion that qualitative and quantitative research are not really different at all.

[17] Those interested in these topics should peruse the pages of *Psychometrika* or the *Journal of Classification*. Krantz, Luce, Suppes, and Tversky (1971-1990) also explore many of these issues.

[18] Pitkin, of course, describes her methodology as "linguistic" analysis, and quantitative researchers might improve themselves by becoming more familiar with her methods.

An exploration of measurement issues along the lines sketched above would, I believe, benefit both quantitative and qualitative researchers, but a discussion of these matters will be worthwhile even if it only shows qualitative researchers how quantitative work must also grapple with complex measurement problems. Because its authors want to be constructive and want to instruct, *Designing Social Inquiry* invariably tries to show how quantitative notions can improve qualitative research. This is laudable, but it leads the authors to neglect mentioning the multitude of problems that confront quantitative researchers, and it ignores the extent to which quantification is based upon qualitative judgments. Both qualitative and quantitative researchers might benefit from a less didactic approach that revealed problems as well as putative solutions. This might lead to a common effort to solve problems of concept formation and measurement that vex both quantitative and qualitative researchers.

Conclusions

Designing Social Inquiry is an excellent sermon, without much condescension, on what qualitative researchers can learn from quantitative researchers. As a work on methodology, it has some substantial defects such as equating explanation with causal inference, proposing a narrow definition of causality, and drawing far too little sustenance from a strong literature on measurement and concept formation. But it also has some substantial strengths. First and foremost, it opens a conversation between qualitative and quantitative researchers, and I think that is very good. Second, its presentation of causal thinking as couterfactual thinking forces researchers to think more carefully about the counterfactuals behind their putative causal models. Third, it has an interesting discussion of selection bias that should be useful to many researchers.[19] Fourth, the final chapter on "Increasing the Number of Observations" is one of the most important notions in the book. I wish KKV had given more concrete examples of how to do this, and I wish they had warned of the dangers of spatial and temporal autocorrelation that can thwart innovative attempts to increase observations, but the basic concept is a very important one.

Students will definitely profit from reading this book. The discipline has already benefitted from the discussions it has kicked off. I look forward to seeing a generation of graduate students uplifted and improved by reciting its useful and informative homilies.

[19] I wish, however, that they had not used the term "selection bias" on page 126 in an example that clearly describes sampling error. The example is described in a section entitled "The Limits of Random Selection" so the authors may have not meant to use the term "selection bias" except in a colloquial fashion, but it is disconcerting, and certainly confusing, nevertheless.

References

Achinstein, Peter (1983), The Nature of Explanation, Oxford University Press: New York.

Andrich, David, (1988), Rasch Models For Measurement, Sage Publications, Inc., Newbury Park.

Bollen, Ken and Richard Lennox, (1991), "Conventional Wisdom on Measurement: A Structural Equation Perspective," Psychological Bulletin, Vol. 110, No. 2, 305-315.

Bollen, Ken A., (1993), "Liberal Democracy- Validity and Method Factors in Cross- National Measures," American Journal of Political Science, Vol. 37, No. 4, 1207-1230.

Baruch A. Brody and Richard E. Grandy (editors) (1989), Readings in the Philosophy of Science, Second Edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Cain, Bruce, John Ferejohn, and Morris Fiorina, (1987), The Personal Vote Constituency Service and Electoral Independence, Cambridge, Massachusetts.

Campbell, Donald T., (1988), Methodology and Epistemology For Social Science, The University of Chicago Press, Chicago.

Campbell, Donald T. and D.W. Fiske (1959), "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, Volume 56: 81-105.

Campbell, Donald T. and Julian C. Stanley (1966), Experimental and Quasi-Experimental Designs for Research, Rand McNally.

Cook, Thomas D. and Donald T. Campbell (1979), Quasi-Experimentation: Design and Analysis for Field Settings, Chicago: Rand McNally.

Duncan, Otis D., (1984a), Notes on Social Measurement Historical and Critical, Russell Sage Foundation, New York.

Duncan, Otis D., (1984b), "Measurement and Structure: Strategies for the Design and Analysis of Subjective Survey Data", Surveying Subjective Phenomena, Russell Sage Foundation: New York.

Fish, M. Steven, Democracy from Scratch, Princeton University Press, Princeton, New Jersey.

Freedman, David A. (1991), "Statistical Methodology and Shoe Leather," Sociological Methdology, Peter Marsden (editor), Jossey-Bass, San Francisco.

Gelman, Andrew and King, Gary (1994), "A Unified Method of Evaluating Electoral Systems and Redistricting Plans," American Journal of Political Science v38, n2:514-554.

Heckman, James J. (1992), "Randomization and Social Policy Evaluation," in Charles F. Manski and Irwin Garfinkel (editors), Evaluating Welfare and Training Programs, Harvard University Press, Cambridge.

Hempel, Carl G. (1965), Aspects of Scientific Explanation, The Free Press: New York.

Holland, Paul W., (1986), "Statistics and Causal Inference", Journal of the American Statistical Association, Vol. 81, No. 396, 945-960.

Kerlinger, Fred N. (1979), Behavioral research : a conceptual approach, New York: Holt, Rinehart, and Winston.

King, Gary, Robert O. Keohane, and Sidney Verba, (1994), Designing Social Inquiry, Princeton University Press, Princeton, New Jersey.

Krantz, David L., R. Duncan Luce, Patrick Suppes, Amos Tversky, (1971, 1989, 1990), Foundations of Measurement, Volumes 1, II, and III. Academic Press, New York.

Pitkin, Hannah (1967), The Concept of Representation, University of California Press: Berkeley.

Rasch, Georg, (1960), Probabilistic Models for Some Intelligence and Attaiment Tests, Copenhagen, Denmark.

Rubin, Donald B. (1974), "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," Journal of Educational Psychology, Volume 66: pages 688-701.

Rubin, Donald B., (1978), "Bayesian Inference for Causal Effects: The Role of Randomization", The Annals of Statistics, Vol. 6, No. 1, 34-58.

Salmon, Wesley (1984), Scientific Explanation and the Causal Structure of the World, Princeton University Press, Princeton, New Jersey.

Scriven, Michael (1975), "Causation as Explanation," Nous, Volume 9, pages 3-10.

Stevens, S. S. (1946), "On the Theory of Scales of Measurement," Science, 103: 677-680.

Stevens, S.S. (1951), "Mathematics, Measurement, and Psychophysics," In S. S. Stevens (editor), Handbook of Experimental Psychology, John Wiley: New York.

Tintner, Gerhard (1952), Econometrics, John Wiley and Sons: New York.

Methods Training at the University of Rochester

Renée M. Smith

At the University of Rochester, our course offerings in political methodology for Ph.D. students include:

PSC 403	Mathematical Modeling D. Austen-Smith
PSC 404	Intro. to Statistical Methods A. Dick/L. Powell
PSC 405	Multivariate Statistical Methods D. Weimer/T. Bird
PSC 406	Survey Design and Analysis D. Niemi
PSC 505	Advanced Statistical Methods R. Smith
PSC 506	Topics in Methods R. Smith

PSC 406 is a two-credit, eight-week course for those interested in survey research. PSC 506 is an advanced methods class offered every two to three years. PSC 403, PSC 404, and PSC 405 are required courses for all students in our department. These three classes along with PSC 505 are the core courses for students pursuing a field in methods.

PSC 403 is designed to teach Ph.D. students the mathematical tools needed for upper-level formal theory and methods classes. The text used is Binmore's (1982) *Mathematical Analysis*. Students take PSC 403 during the fall semester of their first year.

PSC 404 is designed for both Ph.D. and master's level students who have had no statistical training or who want a review. The course also includes a component on research design based on Campbell and Stanley's (1966) *Experimental and Quasi-Experimental Designs for Research*. Students take PSC 404 during the fall semester of their first year. Discrete case examples are followed by continuous case examples once students finish the review of integral calculus in PSC 403. Because students with strong mathematical and statistical skills tend to self-select into Rochester's doctoral program, many Ph.D. students waive PSC 404.

PSC 405 introduces students to the linear regression model as well as to its limitations when assumptions are violated. Both theory and applications are emphasized, and students are generally required to write a research paper in which they collect and analyze data to test a political science hypothesis. The text often used is Hanushek and Jackson's (1977) *Statistical Methods for Social Scientists* supplemented by Kennedy's (1992) *A Guide to Econometrics*. Students take PSC 405 during the spring semester of their first year. Presently, Ph.D. and master's level students take this course together. We have talked about separating the two tracks of students in the future. In exceptional cases, Ph.D. students may waive PSC 405.

PSC 505 is a course for Ph.D. students. The course addresses three broad topics: models with endogenous regressors, maximum likelihood estimators for binary and limited dependent variables, and models for stationary and nonstationary time series. I would love to spend more time on each topic, but the idea is to give students a solid foundation and then let them learn greater detail in PSC 506. Students take PSC 505 during the first semester of their second year. Required texts are Greene's (1993) *Econometric Analysis* and Maddala's (1992) *Introduction to Econometrics*. Students take in-class midterm and final exams and complete a number of homework assignments.

Each homework assignment contains theoretical and applied exercises. I try to coordinate the theoretical proofs with the exercises students run on instructor-supplied datasets. For instance, in their first assignment I almost always have students run a "true" regression and then a regression in which one of the explanatory variables has been residualized. Then I ask them to explain why certain coefficients are the same using (a) formal proofs and (b) words and Venn diagrams. This type of homework problem reflects my philosophy about teaching methods, which is that I want students who are strong in math to understand the intuition behind the proofs and I want less confident students to see that the math is just a formal statement of intuition they may already have. My experience thus far is that both types of students learn something from and respond positively to these exercises.

Upon completion of PSC 505, students may take a comprehensive exam in methods. Some students also take PSC 506 before their exam. In the first half of PSC 506 we go over ML and time series estimators in greater detail. The latter half of the course focuses on student projects and discussions of the methodological issues that arise in their work. There are also a number of good econometrics courses offered in our economics department, and I encourage students serious about methods to take those courses.

The comprehensive exam in methods is offered every January and June. It consists of two parts. In the first, students are given eight hours to solve a series of problems that test their understanding of concepts from each of the core courses. Students are then assigned a journal article that they take home and about which they are required to write a thorough critique. Students can potentially pass the comprehensive exam in methods at one of three levels. Level I indicates the student is capable of teaching undergraduate methods courses. Level II means that the student is able to teach graduate courses through the linear regression model, and Level III indicates that the student can teach advanced graduate methods. Only a pass or fail gets recorded in the student's permanent record, but the student receives a letter from the methods committee indicating at which level s/he passed and what type of recommendation s/he might reasonably expect from our faculty.

In addition to these formal ways of learning methods, we also have an ad hoc methods group that meets whenever someone volunteers to present a paper. It is a great way for students to get feedback from the faculty about work in progress.

Methods Training at Berkeley

Henry E. Brady
University of California, Berkeley

The methods faculty at the University of California, Berkeley includes Henry Brady, David Collier, Robert Powell, Merrill Shanks, and Laura Stoker. The methods field at Berkeley includes three related areas: research design, statistical methods, and formal modelling. Students may elect to complete this field through the course-work option which emphasizes broad understanding and basic competence in all three areas, or they may elect to complete the field through an examination which requires knowledge of all three subareas combined with mastery of either advanced formal modeling or advanced statistical methods.

Introductory Courses and the Course-Work Option — Students who elect the "course work option" to complete the field of political methodology are required to take courses in formal modeling, statistics, and research design with an average grade between A- and B+ or better. Students typically take at least one course from each of the following areas:

(1) Research Design Courses — There are two major introductory research design courses. Both emphasize the actual design of research and cover the basic logic of inference, the threats to validity (using a Campbell-Stanley-Cook framework), measurement problems, representativeness, and other issues. One course, useful for students in all fields but especially important for Americanists, covers a broad array of experimental, survey, case-study, participant observation, time-series, cross- sectional, and time-series cross-sectional designs. The other course focuses on research in comparative politics and spends more time dealing with problems of "small n" case studies, concept formation, and qualitative comparison.

(2) Formal Modeling Course—The introductory formal modeling course covers expected utility theory, basic social choice theory, spatial models of party competition, and game theory through games of incomplete information. The normal and extensive form, Nash equilibria, backward induction, subgame perfection, and other fundamental concepts are discussed. The goal of the course is to make students familiar with the major models and approaches in formal theory.

(3) Introductory Statistics — The introductory statistics course starts with elementary probability theory and works through a Wonnacott and Wonnacott level text to multiple regression. The emphasis is upon learning basic statistical concepts and becoming aware of applications in political science. The principles of representative surveys and experimental design are explored within a statistical framework. The Bernoulli, binomial, Poisson, and normal distributions are introduced with political science examples, such as the frequency of wars for the Poisson. The classical and Bayesian approaches to inference are discussed in detail. Regression is introduced through a thorough analysis of the bivariate case, and the problems of omitted variables, superfluous variables, and multi-collinearity are discussed for multiple regression. Students are required to undertake data analyses with several different data sets.

Advanced Courses and the Examination in Political Methodology — Students may complete the methods field by an all-day examination as well as by course-work. Those who take the examination are expected to cover the three basic fields and to have specialized in at least one area by taking one or more advanced courses. The examination has two sections. The first section covers all three basic fields and students are expected to answer questions in all three areas. The second section involves advanced material from the fields of formal modeling and advanced statistics. Students must complete one of these advanced sections. The examination emphasizes setting up models of political phenomena, problem solving, facility with data analysis, and the ability to identify fruitful and flawed approaches to research design.

Advanced courses include:

(1) Concept formation in comparative politics — This course examines the debates over basic concepts like democracy and corporatism. Readings from statistics (e.g., Ken Bollen's LISREL models of democracy ratings) and from linguistics (e.g., work by George Lakoff) and other sources forms the basis for this course.

(2) Advanced Cross-Sectional Methods — This course, taught approximately at the level of the textbook by Greene, Econometric Analysis, is a sequel to the introductory statistics course described above. Matrix notation is used throughout, although basic concepts are presented in simple scalar models. This course covers errors in variables, factor analysis, recursive models, simultaneous equation models, LISREL, logit and probit models, and other methods applicable to the analysis of cross-sectional data. This and the following course discuss various methods of estimation including generalized least squares, maximum likelihood, and method of moments and various methods of hypothesis testing including the Wald, Lagrange, and likelihood ratio tests.

(3) Advanced Time Series, Time Series Cross-Sectional, and Event History Methods — This course, taught approximately at the level of Greene as well, is also a sequel to the introductory statistics course. Matrix notation is employed throughout this course although many simple examples are presented in scalar notation. The course begins with classic methods for correcting for AR(1) errors, and then moves through Box-Jenkins ARIMA modeling, unit-roots, co-integration, and error correction models. Then the course turns to time series cross- sectional models where both random and fixed effects models are discussed in detail. The course ends with several weeks on event history models including parametric (e.g., Weibull or exponential) models and the Cox proportional hazards model.

(4) Advanced Formal Modeling— Different topics are covered from year to year in this course, but the subject matter includes rational choice models of legislatures, game theory models of negotiating and bargaining in international relations, spatial models of party competition, and social choice approaches to democratic theory.

(5) Practicum in Survey Research — Using the facilities of the Survey Research Center, this course provides a practicum for students to learn about survey research. Students learn about conceptualizing a survey, designing the sample, writing the questionnaire, fielding the study, and analyzing the resulting data.

Who Takes Courses — Although none of these courses are required, more than half of every class takes one or more methods courses. We have an increasing number of students in comparative politics who are taking these courses, and many students avail themselves of the course-work option. A much smaller number (two or three a year) take the methods examination.

Bottom Line — Our goal has been to develop a coherent set of courses that will allow students to move easily from one to another and to get a strong grounding in basic methodological skills. Courses invariably have homework assignments, computer exercises, and, for the more advanced offerings, paper requirements in which original research is done.

Quantitative Research Methods in International Studies Course Syllabus

Philip A. Schrodt
University of Kansas

Course Description: This course is a practical introduction to contemporary quantitative research methods in international politics. It will emphasize hands-on exercises with the data and some of the techniques most commonly used in international politics, as well as with the problems of research design, modelling and testing hypotheses about international behavior. Topics include practical issues in regression analysis; computer programming, dynamic modeling, event data and content analysis, time series techniques, computational techniques such as neural networks and ID3, and classification methods such as cluster analysis.

Prerequisites: Political Science 707 or equivalent. The course is primarily intended for students who will be doing Ph.D.-level research in international politics but should also be of interest to students studying political methodology.

Texts: None required; you will need to buy a book on Pascal programming for the first section. All readings listed will be available in the Allen Room in the "POLS 950" folder.

Comments on Readings:

1. The "Additional" readings are mostly texts and reference works relevant to the topic, along with some example of applications.

2. The readings over-sample from the works of P. Schrodt. This does not mean these works are the finest available in political science, merely that their mode of presentation tends to be compatible with how I teach. When you prepare your own syllabi, I would expect you to be less dependent on the works of this author.

Evaluation: Course grade will be based on the completion of an assortment of homework exercises; these will be approximately weekly. Attendance will not be taken on the assumption that since you are Ph.D. students, you consider this endeavour a full time job, and therefore will attend class. Significant departures from this pattern will be viewed with utmost disfavor.

Assignments should be done on time: in the academic world some deadlines — notably conference papers — are real and close is not good enough. Get in practice. This class also moves fairly quickly across different topics, and it will be difficult to catch up if you fall behind.

Pedagogy

Political methodology is primarily learned by doing, not reading, and therefore this class will focus on practical exercises and techniques rather than the extensive study of texts. Some of these exercises will be fairly time-consuming and therefore you should not wait until the last minute to start them. Articles presenting original research — there are a number of these in the required readings — can also be expected to take some time to understand, so don't attempt to skim them at the last minute.

COURSE TOPICS, EXERCISES AND READING ASSIGNMENTS

Week 1: Introduction (12 January)
Survey of the class, discussion of the general teaching and research approach

Readings: Review material on basic research design from 706 and 707: go back and actually re-read it, and try to work out how much is really relevant when dealing with international relations rather than survey research. Then for a more cautionary and/or countervailing perspective, read:

Eckstein, Harry. "Case Study and Theory in Political Science" in Fred I. Greenstein and Nelson W. Polsby. 1975. *Handbook of Political Science: Strategies of Inquiry.* Addison-Wesley

Most, Benjamin A. and Harvey Starr, "Foreign Policy Substitution and 'Nice' Laws", chpt 5 in Most and Starr, *Inquiry, Logic and International Politics* (South Carolina University Press, 1989)

Bull, Hedley. International Theory: The Case for the Classical Approach. *World Politics* 18: 361- 378.

Donald N. McClosky. "The Rhetoric of Significance Tests", chapter 9 in McCloskey, The Rhetoric of Economics (University of Wisconsin Press, 1985)

Exercise:

Find an article in international relations or comparative politics which you think is flawed because of the research design; make a copy of same and be prepared to discuss it in class.

Additional:

General remarks on bibliography: The primary journals featuring quantitative and modeling research in international relations — beyond the basic crosstabulation/regression approach — are *International Studies Quarterly, Journal of Conflict Resolution* and *International Interactions.* The *American Journal of Political Science* and *American Political Science Review* are also good sources; *AJPS* publishes very little IR but its Workshop sections provide excellent overviews of newer techniques. Over the past decade *World Politics* has begun to publish some quantitative work, and an assortment of smaller specialized journals — for example *Conflict Management and Peace Science* — feature a variety of articles on formal methods.

George, Alexander L. 1979. Case Studies and Theory Development: The Method of Structured, Focused Comparison in Paul G. Lauren ed. *Diplomacy: New Approaches in History, Theory and Policy.* New York: Free Press, pp. 43–68

Holsti, K.J. 1985. *The Dividing Discipline: Hegemony and Diversity in International Theory.* Boston: Allen and Unwin.

Knorr, Klaus and James Rosenau. 1970. *Contending Approaches to International Politics.* Princeton: Princeton University Press.

Nicholson, Michael. 1989. *Formal Theories of International Relations.* Cambridge: Cambridge University Press. [most recent fairly comprehensive survey of IR models]

Vasquez, John A. 1976. "Statistical Findings in International Politics." *International Studies Quarterly* 20:171–218. [a statistical study of statistical studies]

Vasquez, John A. and Marie T. Henehan. 1992. *The Scientific Study of Peace and War.* Lexington, MA: Lexington Books. [undergraduate text on quantitative methods of studying war, mostly from a COW approach.]

Zinnes, Dina A. 1976. *Contemporary Research in International Relations.* New York: Free Press. [good coverage of first generation of behavioralist models]

Weeks 2–4 (16 January – 2 February)
Computer Programming

Readings:

At the Oread Bookstore, Hastings or wherever, find a book that you like on Pascal (ideally, *Turbo Pascal* if you are using DOS; *Think Pascal* if you are using the Macintosh). Buy it; read it.

Additional:

With the ascendency of the personal computer, the literature on computer programming has grown exponentially. The Oread Bookstore has two aisles worth; Borders Books (Overland Park) has at least twice that. Browse.

The following are a few reference books on algorithms; if you are dealing with a standard operation — sorting, indexing, graphics , minimizing a function, generating random numbers — it is a good idea to check some standard references, since the most efficient algorithms are usually anything but obvious. If you are embarking on a large project, it is also worthwhile to consult some works on software development: the amount of time a given project takes can vary by a factor of 20 or more depending on its planning, so investing time in study will pay off in reduced time debugging.

Press, William H, Brian P. Flannery, Saul Teukolsky and William T. Vetterling. 1986. *Numerical Recipes: The Art of Scientific Computing.* Cambridge: Cambridge University Press. (this is an excellent reference for algorithms used in statistical computing; I've also got a disk that contains the code for all of the sample programs in the event you don't wish to retype them]

Sedgewick, Robert. 1988. *Algorithms.* Reading, MA: Addison-Wesley.

Gonnet, G.H. 1984. *Handbook of Algorithms and Data Structures.* Reading, MA: Addison-Wesley.

Kernighan, Brian and P.J. Plauger. 1981. *Software Tools in Pascal.* Reading, MA: Addison-Wesley.

——— 1978. *Elements of Programming Style.* New York: McGraw-Hill (these books are useful because they discuss how to program well in addition to discussing programming)

Bentley, Jon. 1986. *Programming Pearls.* Reading, MA: Addison-Wesley

——— 1988. *More Programming Pearls.* Reading, MA: Addison-Wesley. (highly recommended by McConnell)

McConnell, Steve. 1993. *Code Complete.* Redmond, WA: Microsoft Press. (this is an encyclopedic [830 pp] reference that discusses both the macro issues of project design and control, and the micro issues of coding and data structures. Good references to quantitative studies of factors affecting software development)

Weinberg, Gerald M. 1971. *The Psychology of Computer Programming.* New York: Van Nostrand Reinhold. (just what it says: human factors in software development)

Brooks, Frederick P. 1974. *The Mythical Man-Month.* Reading, MA: Addison-Wesley. (Case study in the [mis]management a large programming project. The five-second summary: "The fact that one woman can produce a child in nine months does not imply that nine women can produce a child in one month.")

Week 5: (9 February)

Review of basic regression and exploratory data analysis

Readings:

Review everything you've read and forgotten about multiple regression from POLS 706 and 707. The focus of the class will be going over the interpretation of a regression analysis and the various things that can go wrong with it.

WARNING: Past experience with this exercise has demonstrated that the level of retention of this material is quite low; to minimize the level of humiliation, review it.

King, Gary. 1986. How Not to Lie with Statistics. *American Journal of Political Science* 30,3: 666–687.

Additional:

Diaconis, Persi and Bradley Efron. 1983. "Computer-Intensive Methods in Statistics." *Scientific American.* 248,5:116–130 [nice description of bootstrapping]

Berry, William D. and Stanley Feldman. 1985. *Multiple Regression in Practice.* Beverly Hills: Sage.

Everitt, B.S. and G. Dunn. 1983. *Advanced Methods of Data Exploration and Modeling.* London: Heinemann Books

Freedman, David Robert Pisani and Rogers Purves. 1978. *Statistics.* New York: Norton (this is the archetypical, and controversial, "new wave" statistics textbook — lots of pictures, no equations)

Mooney, Christopher Z. and Robert D. Duval. 1993. *Bootstrapping: A Nonparametric Approach to Statistical Inference.* Newbury Park: Sage.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press

Tukey, John W.. 1977. *Exploratory Data Analysis* Reading, MA: Addison-Wesley.

Week 6 (16 February)

Univariate Dynamic Models and Chaos

Readings

Schrodt, "Dynamic Models" chapter

Richards, Diana. 1993. "A Chaotic Model of Power Concentration in the International System." *International Studies Quarterly* 37:55–72.

Additional

General references on difference equations

Goldberg, Samuel. 1958, 1986. *Introduction to Difference Equations.* New York: Wiley (1986 edition by Dover)

Huckfeldt, R. Robert, C.W. Kohfeld, and Thomas W. Likens. 1982. *Dynamic Modelling: An Introduction.* Beverly Hills: Sage.

Kelley, Walter G. and Allan C. Peterson. 1991. *Difference Equations: An Introduction with Applications.* Boston: Academic Press.

Lakshmikantham, V. and D. Trigiante. 1988. *Theory of Difference Equations: Numerical Methods and Applications.* New York: Academic Press.

Luenberger, David G.. 1979. *Introduction to Dynamic Systems: Theory, Models and Applications.* New York: Wiley. [good introduction to applications of differential equation models]

Mickens, Ronald E. 1990. *Difference Equations: Theory and Applications.* New York: Van Nostrand Reinhold.

Olinick, Michael. 1978. *An Introduction to Mathematical Models in the Social and Life Sciences.* Reading, Mass: Addison-Wesley.

Chaos:

Barton, Scott. 1994. "Chaos, Self-Organization and Psychology." *American Psychologist* 49: 5–14. [short discussion of the advantages and problems of trying to apply chaos to human behavior]

Casti, John L. 1989. *Alternative Realities: Mathematical Models of Nature and Man.* New York: Wiley. [excellent coverage of both chaos and evolutionary models, though the mathematics gets a little heavy at times. The discussions are readable though.]

Devaney, Robert L. 1986. *An Introduction to Chaotic Dynamical Systems.* Menlo Park: Cummings. [fairly heavy mathematics...]

Gleick, James. 1987. *Chaos: Making a New Science.* New York: Viking.[how to write a best-seller about mathematical modelling...]

Hill, Walter W. 1992. "Deterministic Quasi-Periodic Behavior of an Arms Race Model." *Conflict Management and Peace Science* 12:79–98.

Richards, Diana. 1992. "Spatial Correlation Test for Chaotic Dynamics in Political Science." *American Journal of Political Science* 36:1047–1069

Saperstein, Alvin M. and Gottfried Mayer-Kress. 1988. "A Nonlinear Model of the Impact of SDI on the Arms Race." *Journal of Conflict Resolution* 32,4: 636–670. [application of chaos theory to the arms race problem]

Saperstein, Alvin M. 1992. "Alliance Building versus Independent Action: A Nonlinear Modeling Approach to Comparative International Stability" *Journal of Conflict Resolution* 36: 518–545. [application of chaos theory to crisis stability]

Week 7 (23 February)

Dynamic Models: Systems

Readings

Schrodt, "Dynamic Systems" chapter.

Francisco, Ronald A. 1993. "The Relationship Between Coercion and Protest: An Empirical Evaluation in Three Coercive State" (manuscript)

Optional Review:

J. Johnston. 1972. *Econometrics.* New York: McGraw-Hill. Chapter 4 (linear algebra)

Additional:

Note: The general books on difference equations cited above also deal with systems.

Intriligator, Michael D. and Dagobert L. Brito. 1989. "Richardsonian Arms Race Models" in Manus Midlarsky, ed. *Handbook of War Studies.* Boston: Unwin Hyman. [Emphases the rational variants on the model]

Luterbacher, Urs and Michael Don Ward (eds.). 1985. *Dynamic Models of International Conflict.* Boulder,CO: Lynne Rienner Publishing.

May, Robert M. 1974. *Stability and Complexity in Model Ecosystems.* Princeton: Princeton University Press.

McGinnis, Michael D. 1991. "Richardson, Rationality and Restrictive Models of Arms Races." *Journal of Conflict Resolution* 35:443–473. [Extensive bibliography. The literature on the Richardson model is huge — a couple hundred articles — and I've got some other bibliographies if you are interested.]

Mintz, Alex and Michael D. Ward. 1989. "The Political Economy of Military Spending in Israel." *American Political Science Review* 83:521–533. [nice example of the specification and numerical estimation of a dynamic model]

Rapoport, Anatol. 1974. *Fights, Games and Debates.* Ann Arbor: University of Michigan Press.

Richardson, Lewis F. 1960. *Statistics of Deadly Quarrels.* Chicago: Quadrangle Books.

Richardson, Lewis F. 1960. *Arms and Insecurity.* Chicago: Quadrangle Books.

Ward, Michael D. ed. 1985. *Theories, Models and Simulations in International Relations.* Boulder, CO: Westview Press. [survey of simulation models]

Week 8 (2 March)

Time Series Analysis: Tests of Causality

Readings:

Enders, Walter and Todd Sandler. 1993. "The Effectiveness of Antiterrorism Policies: A Vector- Autoregression-Intervention Analysis." *American Political Science Review* 87:829–844.

Freeman, John. 1983. "Granger Causality and Time Series Analysis of Political Relationships." *American Journal of Political Science.* 27,2:327–358

Freeman, John R, John T. Williams and Tse-Min Lin. 1989. Vector Autoregression and the Study of Politics." *American Journal of Political Science.* 33:825–841

Additional:

Beck, Nathaniel. 1983. "Time-Varying Parameter Regression Models" *American Journal of Political Science.* 27:557–600. [discussion of various tests for determining whether a change has occurred in the parameters of a time series model]

Davis, David R. and Michael D. Ward. 1990. "They Dance Alone: Deaths and the Disappeared in Contemporary Chile." *Journal of Conflict Resolution* 34: 449–475. [VAR/event data analysis of causes and effects of repressive activities in Chile].

Goldstein, J. S., and John. R. Freeman (1990) *Three-Way Street: Strategic Reciprocity in World Politics.* Chicago: University of Chicago Press. [extensive VAR analysis of event data]

McCleary, Richard and Richard A. Hay Jr. 1980. *Applied Time Series Analysis for the Social Sciences.* Beverly Hills: Sage. [one of the more readable expositions on quasi-experimental design and Box-Jenkins]

McDowall,David D., Richard McCleary, Errol E. Meidinger and Richard A. Hay. 1980. *Interupted Time Series Analysis.* Beverly Hills: Sage [this also has a good bibliography]

McGinnis, Michael D. and John T. Williams. 1989. "Change and Stability in Superpower Rivalry" *American Political Science Review* 83:1101–1123. [VAR analysis using COPDAB, military expenditures; similar in approach to Goldstein and Freeman]

Sheehan ,Richard G. and Robin Grieves. 1982. "Sunspots and Cycles: A Test of Causation." *Southern Economic Journal* (January 1982):775–777. [statistical demonstration that business cycles cause sunspots]

Thurman, Walter N. and Mark E. Fisher. 1988. Chickens, Eggs and Causality, or Which Came First? *American Journal of Agricultural Economics* 70(2):237–238.[statistical demonstration that eggs cause chickens but not vice versa]

General books on time series [this is just a sample]

T.W. Anderson, T.W.. 1972. *The Statistical Analysis of Time* Series. New York: Wiley. [very nice mathematical treatment of the basic processes]

Box, George E.P. and G.M. Jenkins. 1970. *Time-series Analysis: Forecasting and Control.* San Francisco: Holden-Day [this is the original exposition of the Box-Jenkins-Tiao techniques; it is not particularly easy to read and more recent treatments would be better]

Chatfield, C. 1989. *The Analysis of Time Series: An Introduction.* London: Chapman and Hall.

Gottman, John M.. 1981. *Time-Series Analysis: A Comprehensive Introduction for Social Scientists.* Cambridge: Cambridge University Press

Granger, C.W.J. and Paul Newbold. 1986. *Forecasting Economic Time Series.* San Diego: Academic Press.

Harvey, Andrew C. 1993. *Time Series Models.* Cambridge: MIT Press.

Kendall, M.G.. 1973. *Time-Series.* New York: Hafner Press [excellent, if terse, practical introduction to the basic techniques]

Ostrom, Charles W.. 1978. *Time Series Analysis: Regression Techniques.* Beverly Hills: Sage [deals primarily with autocorrelated error but has some discussion of correlograms]

Nelson, Charles R. and Leejoon Kang. 1984. "Pitfalls in the Use of Time as an Explanatory Variable in Regression." *Journal of Business and Economic Statistics.* 2,1:73–82. (CC)

Stimson, James A.. 1985. "Regression in Time and Space: A Statistical Essay." *American Journal of Political Science* 29,4:914–947.

Week 9 (9 March)

Time Series Analysis: Stochastic Models

Readings:

Batschelet, Edward . 1976. *Introduction to Mathematics for Life Scientists.* New York: Springer- Verlag. pp. 446– 463 (try to read the material on the Poisson carefully, then skim the material on the normal and other distributions)

Berry, Frances Stokes and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84:395–415. [okay, okay, it isn't IR but it is a good example of how to use the method]

Beck, Nathaniel 1991. "The Illusion of Cycles in International Relations." *International Studies Quarterly* 35,4: 455–476. [also reply by Goldstein]

Geller, Daniel S. 1993. "Power Differentials and War in Rival Dyads." *International Studies Quarterly* 37:173–193. [Markov chain analysis]

King, Gary. 1989. "Event Count Models for International Relations: Generalizations and Applications" *International Studies Quarterly* 33(2):123–148.

Additional:

Allison, Paul D.. 1984. *Event History Analysis* Beverly Hills: Sage

Bartholomew, J.D. 1973. *Stochastic Models of Social Processes.* New York: Wiley. [good general introduction to stochastic models; most of the examples are from sociology]

Casstevens, Thomas W. 1989. "The Circulation of Elites: A Review and Critique of a Class of Models." *American Journal of Political Science* 33: 294–317. [covers a lot of the Poisson stuff]

Houweling, Henk W. and J. B. Kune. 1984. "Do Outbreaks of War Follow a Poisson-Process?" it Journal of Conflict Resolution 28,1: 51–62.

Houweling, Henk W. and Jan G. Siccama. 1985. "The Epidemeology of War, 1816–1980." *Journal of Conflict Resolution* 29,4: 641–664.

King, Gary. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* Cambridge: Cambridge University Press. [maximum likelihood methods applied to various models of political behavior]

Midlarsky, Manus I. 1986. "A Hierarchical Equilibrium Theory of Systemic War." *International Studies Quarterly* 30,1:77–106

Richardson, Lewis F. 1960b. *Statistics of Deadly Quarrels.* Pittsburgh, PA: Boxwood Press. [original Poisson test on war distribution]

Schrodt, Philip A. 1985. The Role of Stochastic Models in International Relations Research. pp. 199–221 in Michael Don Ward (ed.) *Theories, Models and Simulation in International Relations.* Boulder, CO: Westview.

Week 10 (16 March)

KEDS (Kansas Event Data System) and Event Data Analysis

Readings:

Schrodt, Philip A. 1994. "Event Data in Foreign Policy Analysis" in Patrick J. Haney, Laura Neack, and Jeanne

A.K. Hey. *Foreign Policy Analysis: Continuity and Change.* New York: Prentice-Hall

Schrodt, Philip A., Shannon Davis and Judith L. Weddle. 1995. "KEDS: A Program for the Machine Coding of Event Data." forthcoming, *Social Science Computer Review* (Winter, 1995).

Laurence, Edward J. 1990. "Events Data and Policy Analysis." *Policy Sciences* 23:111–132.

Leng, Russell J. and J. David Singer. 1988. "Militarized Interactive Crisises: The BCOW Typology and Its Applications." *International Studies Quarterly* 32:155–1174.

Additional:

Andriole, Stephen J. and Gerald W. Hopple. 1984. "The Rise and Fall of Events Data: From Basic Research to Applied Use in the U.S. Department of Defense." *International Interactions* 11:293– 309.

Azar, Edward E. 1982. *The Codebook of the Conflict and Peace Data Bank (COPDAB).* College Park, MD: Center for International Development, University of Maryland.

Azar, Edward E., Richard A. Brody and Charles A. McClelland (eds.) 1972. *International Events Interaction Analysis: Some Research Considerations.* Beverly Hills: Sage Publications.

Azar, Edward E. and Joseph Ben-Dak. 1975. *Theory and Practice of Events Research.* New York: Gordon and Breach.

Burgess, Philip M. and Raymond W. Lawton. 1972. *Indicators of International Behavior: An Assessment of Events Data Research.* Beverly Hills: Sage Publications.

Daly, Judith Ayres and Stephen J. Andriole. 1980. "The Use of Events/Interaction Research by the Intelligence Community." *Policy Sciences* 12:215–236.

Gerner, Deborah J., Philip A. Schrodt, Ronald A. Francisco, and Judith L. Weddle. 1994. The Machine Coding of Events from Regional and International Sources, *International Studies Quarterly* 38:91–119.

Heise, David. 1988a. "Modeling Event Structures." *Journal of Mathematical Sociology* 13:138–168.

Heise, David. 1988b. "Computer Analysis of Cultural Structures." *Social Science Computer Review* 6:183–196.

Hermann, Charles, Maurice A. East, Margaret G. Hermann, Barbara G. Salmore, and Stephen A. Salmore. 1973. *CREON: A Foreign Events Data Set.* Beverly Hills: Sage Publications.

Hoople, Gerald W., Stephen J. Andriole and Amos Freedy (eds.). 1984. *National Security Crisis Forecasting and Management.* Boulder: Westview Press.

International Studies Quarterly. 1983. "Symposium: Event data Collections." *International Studies Quarterly* 27.

Lehnert, Wendy and B. Sundheim. 1991 "A Performance Evaluation of Text Analysis." *AI Magazine* 12:81–94.

McClelland, Charles A. 1961. "The Acute International Crisis." *World Politics* 14:184–204.

McGowan, Patrick, Harvey Starr, Gretchen Hower, Richard L. Merritt and Dina A. Zinnes. 1988. "International Data as a National Resource." *International Interactions* 14:101–113.

Merritt, Richard L., Robert G. Muncaster and Dina A. Zinnes (eds.) 1993. *International Event Data Developments: DDIR Phase II.* Ann Arbor: University of Michigan Press.

Munton, Donald. 1978. *Measuring International Behavior: Public Sources, Events and Validity.* Dalhousie University: Centre for Foreign Policy Studies.

Peterson, Sophia. 1975. "Research on research: Events data studies, 1961–1972." In Patrick J. McGowan (ed.) *Sage International Yearbook on Foreign Policy Studies,* 3. Beverly Hills: Sage Publications.

Salton, G. (1989) *Automatic Text Processing.* Reading, Mass: Addison-Wesley.

Schrodt, Philip A. and Alex Mintz. 1988. "A Conditional Probability Analysis of Regional Interactions in the Middle East." *American Journal of Political Science* 32: 217–230.

Schrodt, Philip A., and Deborah J. Gerner. 1994. "Statistical Patterns in a Dense Event Data Set for the Middle East, 1982–1992." *American Journal of Political Science* 38.

Sigler, John H., John O. Field and Murray L. Adelman. 1972. *Applications of Event Data Analysis: Cases, Issues and Programs in International Interaction.* Beverly Hills: Sage Publications.

Vincent, Jack E. 1983. "WEIS vs. COPDAB: Correspondence Problems." *International Studies Quarterly* 27:160–169.

Weber, Robert Philip. 1990. *Basic Content Analysis.*(2nd ed) Beverly Hills: Sage.

—— SPRING AND ISA BREAK ——

Week 11: (6 April)

Cluster analysis and nearest-neighbor techniques

Readings:

Everitt, Brian. 1980. *Cluster Analysis* (2nd ed.) New York:Wiley/Halsted. pp. 1–58 [just read for the concepts; don't try to get all of the math]

Schrodt, *Patterns, Rules and Learning,* chapter 5, section on nearest neighbor methods

Additional:

Aldenderfer, Mark S. and Roger K. Blashfield. 1984. *Cluster Analysis.* Newbury Park: Sage.

Aldrich, John and Charles F. Cnudde. 1975. "Probing the Bounds of Conventional Wisdom: A Comparison of Regression, Probit and Discriminant Analysis." *American Journal of Political Science* 19,3:571–608. [comparison of these methods; shows that there isn't a whole heck of a lot of difference]

Anderson, E.B.. 1980. *Discrete Statistical Models with Social Science Applications.* Amsterdam: Elsevier/North Holland.

Dunteman, George H. 1989. *Principle Components Analysis.* Newbury Park, CA: Sage.

Feinberg, S.E.. 1977. *The Analysis of Cross-Classified Categorical Data.* Cambridge: MIT Press.

Greenacre, Michael J. 1984. *Theory and Applications of Correspondence Analysis.* New York: Academic Press.

Klecka, William. 1978. *Discriminant Analysis.* Beverly Hills: Sage

Malhotra, Naresh K.. 1983. "A Comparison of the Predictive Validity of Procedures for Analyzing Binary Data." *Journal of Business and Economic Statistics* 1,4:326–336 [another comparison of regression, discriminant, logit and probit — again, not a lot of difference]

Schalkoff, Robert. 1992. *Pattern Recognition: Statistical, Structural and Neural Approaches.* New York: Wiley. [nice combination of statistical and computational approaches]

Week 12: (13 April)

Rule-based models: Expert Systems and ID3

Readings:

Majeski, Stephen. 1989. "A Ruled-Based Model of the United States Military Expenditure Decision- Making Process." *International Interactions* 15: 129–154.

Schrodt, *Patterns, Rules and Learning,* chapter 4, chapter 5 section on ID3

Stewart, Philip D., Margaret G. Hermann and Charles F. Hermann. 1989. "Modeling the 1973 Soviet Decision to Support Egypt." *American Political Science Review* 83: 35–60.

Taber, Charles S. 1992. "POLI: An Expert System Model of U.S. Foreign Policy Belief Systems." *American Political Science Review* 86,4:888–904.

Additional:

Andriole, Stephen J.and Gerald W. Hopple. 1988. *Defense Applications of Artificial Intelligence.* Lexington MA: Lexington.

Robert A. Benfer, Edward E. Brent and Louanna Furbee. 1991. *Expert Systems* Beverly Hills: Sage

Brent, Edward. 1988. "New Approaches to Expert Systems and Artificial Intelligence Programming." *Social Science Computer Review* 6(4):569–578

Cimbala, Stephen. 1987. *Artificial Intelligence and National Security.* Lexington, MA: Lexington Books.

Hudson, Valerie, ed. 1991. *Artificial Intelligence and International Politics.* Boulder: Westview [comprehensive collection of the major AI/IR projects]

Levine, Robert I., Diane E. Drang and Barry Edelson. 1990. *AI and Expert Systems.* New York: McGraw Hill [AI without LISP: provides basic algorithms in Pascal]

Sylvan, Donald A. and Steve Chan, 1984. *Foreign Policy Decision Making: Perception, Cognition and Artificial Intelligence.* New York: Praeger.

Winston, Patrick Henry. 1984. *Artificial Intelligence 2nd edition.* Reading, MA: Addison-Wesley. [standard AI textbook; it's into a 3rd or 4th edition now]

Week 13: (20 April)

Genetic algorithms and neural networks

Readings:

Axelrod, Robert. 1987. "The Evolution of Strategies in the Iterated Prisoners' Dilemma" pp 32–41 in Lawrence Davis, ed. *Genetic Algorithms and Simulated Annealing.* Los Altos,CA: Morgan Kaufmann.

Schrodt, *Patterns, Rules and Learning,* chapter 5, sections on neural networks and genetic algorithms

Richard Kimber, "Artificial Intelligence and the Study of Democracy" *Social Science Computer Review* 9

G. David Garson, "A Comparison of Neural Network and Expert Systems Algorithms with Common Multivariate Procedures for Analysis of Social Science Data" *Social Science Computer Review* 9

Additional:

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books. [first well-known application of evolutionary techniques to IR]

Freeman, James A. and David M. Skapura. 1991. *Neural Networks: Algorithms, Applications and Programming Techniques.* Reading. MA: Addison-Wesley.

Goldberg, David E. 1989. *Genetic Algorithms in Search, Optimizations and Machine Learning.* Reading, Mass: Addison-Wesley. [text on GA's]

Hertz, J., A. Krough and R.G. Palmer. 1990. *Introduction to the Theory of Neural Computing.* Reading, MA: Addison-Wesley.

Holland, John H. 1975. *Adaptation in Natural and Artificial Systems.* Ann Arbor: University of Michigan Press. [original Holland; most of the recent work has gone in somewhat different directions]

Kollman, Ken, John H. Miller and Scott E. Page. 1992. "Adaptive Parties in Spatial Elections." *American Political Science Review* 86,4:929–937. [applies genetic algorithms, among other methods, to the problem of optimal party position in a spatial model]

Rumelhart, David E., James L. McClelland and the PDP Research Group. 1986. *Parallel Distributed Processing* (2 vols). Cambridge: MIT Press. [more than everything you ever wanted to know about neural network methods]

Wasserman, Philip D. 1989. *Neural Computing: Theory and Practice.* New York: Van Nostrand Reinhold. [textbook: considerably more readable than Rumelhart et al]

Weiss, Sholom M. and Casimir A. Kulikowski. 1990. *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems.* San Mateo CA: Morgan Kaufmann.

Week 14: (27 April)

Mathematica and statistical graphics.

Demonstration:

Philip Huxtable will demonstrate some simulation and analytical methods using Mathematica

Readings:

Wilkinson, Leland. 1989. "Cognitive science and graphic design" chapter in *SYGRAPH: The System for Graphics.* Evanston, IL: SYSTAT,Inc.

Additional:

Wilkinson has an extensive bibliography of articles; the key thing to keep in mind is that good statistical graphics are carefully designed based on the characteristics of human visual perception; they don't simply arise by luck (bad statistical graphics can also be carefully designed: consult any issue of USA Today or Time). You might also find it useful to browse a couple of books on basic graphical design; the bookstore has dozens of them in the art section.

Cleveland, W.S. 1985. *The Elements of Graphing Data.* Monterey,CA: Wadsworth.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

Tukey, John W. 1977. *Exploratory Data Analysis.* Reading, MA: Addison-Wesley.

Game-Theoretic and Empirical Methodologies: *Ever* the Two Shall Meet?[1]

Scott Gates
Michigan State University and the University of Trondheim, Norway

Sherry Bennett Quiñones
Rice University

Introduction

The call for a modeling dialogue between proponents of rational choice models and advocates of empirical approaches to political science continues to grow.[2] However, to suggest that proponents from each modeling tradition simply fail to communicate is specious, if not misleading and altogether incorrect. Indeed, numerous examples of rational choice models that have been empirically tested abound.[3] Although most will agree that utilizing both conventions is not necessarily a novel undertaking, few attempts have been made to explicitly discuss paths of conceptualization and operationalization for these two methodologies. As Myerson notes, "no analytical discipline can solve real world problems unless it is complemented by a modeling dialogue... a process in which theorists and empiricists work together interactively on the difficult task of finding tractable models that capture and clarify the important aspects of real situations" (1992: 63-64). Our purpose in this essay is to impart some thoughts concerning the inherent differences and commonalities in these formal approaches, in order to demonstrate the utility of integrating both methodologies. Both provide rich and rigorous paths to understanding complex social phenomena. Each approach is used to some extent to answer different questions. However, this should not negate the utility of a "modeling dialogue." In some instances opportunities in research arise, which are conducive to multiple methods of analysis. Thus, we argue that integrating both analytical traditions could be especially fruitful.[4]

[1] We thank Sara McLaughlin, Cliff Morgan, Ido Oren, and Rick Wilson for their useful suggestions. Of course, we are the ones responsible for what is written here; none of those who provided comments should be implicated.

[2] Rational choice, in a general sense, embodies various literatures including social choice, public choice, game theory, and rational actor models. Herein, most of our comments pertain specifically to models of the game theoretic variety.

[3] We discuss a small number of these works later in this essay. Generally speaking we focus our attention on the International Relations literature. For an extensive discussion of prominent works in the American Politics literature see Green and Shapiro (1994). Also see Lalman et al. (1993) for works within both fields.

[4] At some length Green and Shapiro (1994) make a similar argument. However, unlike Green and Shapiro (1994) our purpose

Recently, a small but growing body of research in political science has attempted to integrate the two approaches. These examples link the two approaches indirectly, whereby a game-theoretic model is used to derive a set of propositions which in turn are empirically analyzed.[5] The limited integration of both the game-theoretic and empirical statistical traditions ultimately stems from the complexities associated with combining both approaches. Foremost in complicating the integration of the two methodologies, is the inherent makeup of the two traditions. By their very nature, "empirical and formal analytic traditions possess an indeterminacy of translation" (Brehm and Gates 1992: 3). Putnam (1985) illustrates the complexity of this issue as he comments "If any problem has emerged as the problem for analytic philosophy in the twentieth century, it is the problem of how words 'hook onto' the world."

Numerous practical problems surface when attempting to test game theoretic models empirically. One of the central problems is how to operationalize game theoretic constructs into variables that can be used in an empirical test. Throughout this essay, the argument will be made that the problems which surface in both of these modeling traditions are common, if not insurmountable.[6] We contend, however, that the combination of both modeling traditions can only serve to elevate the validity and rigor in which models are specified. To consider the issues associated with combining these two methodologies, it is first necessary to discuss the modeling goals and orientation behind each tradition.

Contrasting Game Theoretic and Empirical Methodologies

Quantitative and formal approaches to social science both seek to explain, predict, and understand human behavior. As such, it is not so much the goals that differ, but rather the orientations of each which complicate uniting both methodologies. The basic orientation of game theoretic analysis is to focus on a particular aspect of a social interaction. Such formal models allow us to vary clear and explicit assumptions a tiny bit and examine the implications through formal analysis. We learn from the accumulation of models. No single model explains everything. In such a way game theoretic models provide a progressive understanding of the world that is different but not at odds with the empirical modeling approach. To set out to directly "test" game theoretic models misses the point. There is considerable room for integration whereby propositions derived from formal models can be examined empirically.

Analyzing each approach in terms of how they proceed with the "hypothetico-deductive"[7] ("scientific") method of research, helps shed light on the differences and similarities of both approaches. This enables us to construct an argument as to how the integration of both approaches, although complicated, can be achieved.

Conceptualization

When deducing a theory to explain a particular research question, the process of conceptualization in forming a model is extremely important. In fact, a model is one of the most integral parts of the "hypothetico-deductive" method. As a result, it is also one of the most complicated issues associated with the integration of game theory and empirical modeling. Conceptualization involves specification. In empirical models, a theory is used to enumerate fundamental concepts for a model; the end result is the specification of a set of indicators which indicate the presence or absence of a concept under study. In game-theoretic analysis, conceptualization involves the deduction of the formal game structure, including: the players, the structure of the payoffs, assignment of decision nodes to players, actions, information sets,[8] and probability distributions for each node (Kreps 1990: 355-384).

Game theory has been described as "a special conceptual structure for organizing and structuring thoughts about concepts in an attempt to order them and predict their effects" (Myerson 1992: 62). Game theoretic formalizations (i.e. the specification of actions, payoffs, information, etc.) naturally impose restrictions on a model. "Rational choice theorists deliberately simplify and abstract reality in their models. Game models do not even attempt to address all the complexity of the social world. Instead, they focus on certain elements of social situations to lay bare how motivations and actions are interrelated" (Morrow 1994: 8).

is not to demonstrate the empirical irregularities associated with rational choice work. On the contrary, it is our intention to demonstrate how rational choice models can be employed to vigorously specify empirical models and to elicit insights among competing empirical models. A similar argument can be made for the utility of empirical models eliciting support for competing rational choice models. However, we do not elaborate upon this issue. See Wilson and Herzberg (1994) for a discussion of this point.

[5]Other formal deductive methods are more conducive to direct statistical evaluation. Dow and Hinich (1992) link formal spatial models and empirical statistical models more directly. Gerber and Jackson (1993) and Achen (1992) even more directly link formal Bayesian models to statistical analyses. Their work comes the closest to what Blackwell and Girshick (1954) set up in their *Theory of Games and Statistical Decisions* where a more direct link is established between game theoretic and statistical models. What we are concerned with here is the more indirect approach; we are not advocating the direct testing of game-theoretic models or rational choice assumptions in general.

[6]Note Quine's (1953) critique of the analytic-synthetic distinction.

[7]See King et al (1994).

[8]Technically a player i's information set ω_i at any particular point of a game is the set of different nodes in the game tree that she knows might be the actual node, but between which she cannot distinguish by direct observation.

The enterprise of game theoretic modeling is oriented towards clarification and simplification of complex social interactions. The goal is to provide an understanding of specific aspects of different social phenomenon, not to provide a complete picture.

Ceteris paribus conditions, for example, provide the restrictions that are imposed on any game theoretic model. These restrictions represent initial conditions within a model and the parameters that will be manipulated. Although restrictions in a game-theoretic model may be justified analytically, such simplifications could cause problems when attempting to specify the model empirically. Empirical specifications of game-theoretic models might prove to be underspecified because they fail to incorporate enough parameters into a model. This point becomes particularly acute in the context of game theory when a model produces no equilibrium or multiple equilibria. In such cases, "The modeler has failed to provide a full and precise prediction of what will happen" (Rasmusen 1989: 27). Game theorists have devoted considerable energy towards devising new concepts of equilibria refinement to explain such indeterminacy as seen in the Folk Theorem or when off-the-path or out-of-equilibrium behavior occurs.

Whereas conceptualization appears to be a strictly delineated activity in game theoretic modeling, to some degree it is less precise in empirical research. Each step in a game analytic model is deduced in a meticulous manner. Empirical models can lack analytic rigor and sometimes run into overspecification problems. Game theoretic models, on the other hand, possess analytic rigor, but at the expense of underspecification. What does this mean about integrating the two approaches? To answer this question, it necessary to discuss the similarities and differences between the two approaches with respect to the process of operationalization. An analysis of the process of operationalization suggests a different orientation between the two methodologies.

Operationalization

The next step in the hypothetico-deductive method involves extracting a set of testable hypotheses from a model. The process of operationalization is vital in this step and also highlights differences between the two modeling traditions. In empirical models, the process of operationalization involves taking abstract constructs from a theory and determining what variables can be used to represent them. These variables are then ordered according to their relationship in a causal manner. This subsequently allows the empirical modeler to select indicators which can be used to measure variables. The process of operationalization for game theoretic models, on the other hand, is not so straight forward.

In game theoretic models, operationalization involves a process of delineating strategies which produce equilibria. Having initially conceptualized a given theory into a game theoretic form, the next stage begins the analysis of the strategies that produce solutions for the game. Equilibria can take different forms: Nash, separating, pooling, correlated, etc. Refinements of equilibrium concepts usually involve the identification of *out-of-equilibrium* or *off-the-path* information sets for players in a model. Out-of-equilibrium actions can be viewed as "tests" of out-of-equilibrium responses (Kreps 1990: 417-449). In game theoretic models it is believed that players condition their actions on what they believe is "rational" under certain circumstances. As such, they are "pretty sure" they know what will happen if they deviate from alleged rational action (Kreps 1990: 387).[9] With respect to operationalization, out-of-equilibrium behavior and refinements of Nash equilibrium generally facilitate the delineation of actions which would appear to contradict a theory. "As in any analytic approach to real-life problems, the best we can hope for is to have a class of models sufficiently rich and flexible that, if anyone objects that our model has neglected some important aspect of the situation that we are trying to analyze, we can generate a more complicated extension of our model that takes this aspect into account" (Myerson 1991: 83).

When comparing the process of operationalization within each modeling framework, one of the main weaknesses apparent in game theoretic models is their inability to translate opaque constructs into tangible variables. While operationalizing constructs into variables within an empirical framework is often problematic, it is even a greater issue for game theoretic models. In fact, this is perhaps one of the main weaknesses associated with integrating the two approaches. For example, when thinking about the notion of a *signal* in a game theoretic model, how can this concept be broken down into something tangible that can be observed? What specific actions compose a signal? Moreover, how would one operationalize a trigger strategy? What actions form a trigger? Answers to these questions are not straight forward, serving to complicate the integration of the two modeling approaches.

Interpretation

The final steps in the hypothetico-deductive method pertain to analyzing the results derived from a model and subsequently interpreting them. Again, this process varies among the two frameworks. Empirical models generally are set up to "confirm" or "disconfirm" testable hypotheses by examining results derived from statistical analysis. Depending upon the type of analysis conducted, coefficient estimates are typically analyzed with respect to their direction and magnitude on a posited relationship and in terms of their statistical significance. Depending upon the model's

[9] Recall the two central assumptions of game theory, utility maximization and common knowledge.

posited relationships and its respective statistical results, the viability of a theory is then assessed.

Game theoretic models are not characterized by these same types of problems. Although propositions are similarly derived from the model, game theoretic models do not set up testable propositions in the same manner. Analysis consists of setting up propositions and then illustrating how they can be confirmed. Refinements of Nash equilibria to some degree illustrate the revision of a theory. However, such refinements usually consist of explaining deviations or altering a formalization in a manner which explains out-of-equilibrium behavior. The problem is not usually associated with the "solution concept," but with the game theoretic model itself (Kreps 1990: 417-449). If actions take place that may seem unexpected, then it is necessary to include those actions within the game theoretic model. Under such conditions, the game theoretic model can be considered to be underspecified (Johnson 1993: 79-80). Game theoretic models are rarely explicitly tested with actual observations in the real world. Experiments are the most widely used empirical approach used to test game theoretic models.[10] We suggest that the rigorous manner in which game theoretic models are deduced, combined with an empirical framework that permits it to be tested in a real world context, can allow political scientists to evaluate theories in a more rigid manner.

Although the differences between the two approaches seem to make their integration an arduous process, these complications are nonetheless surmountable. This is especially true when taking into account the similarities that exist between the two approaches. In fact, a focus on the similarities which exist between the two approaches helps illustrate how valuable game models can be in empirically specifying models.

Similarities Between Game Theoretic and Empirical Models

Generally, the operations that empirical and game theoretic modelers follow in illustrating beliefs about behavior is very similar. Each approach is involved in extracting information from a theory and incorporating it into a model. Moreover, each imposes restrictions upon models. These usually take the form of initial conditions and parameters which are deemed relevant to the model. However, game theoretic models begin with initial conditions or assumptions that are not specifically integrated into the model itself. These types of restrictions are placed upon the model in an attempt to maintain simplicity and rigor. These restrictions, in turn, may result in underspecification problems if the model were to be empirically tested. How is it possible for game theoretic models to aid in empirical specification if they are underspecified from the beginning? It is possible because both game theoretic and empirical research are engaged in ordering relevant concepts within their respective models.

In the process of operationalization, each method goes through the process of ordering concepts. Likewise each is involved in the temporal ordering of factors which are explicitly modeled as relevant in explaining behavior. Certain causal factors are specified which are believed to explain certain behaviors. Game theoretic models generally describe preferences over outcomes. However, game theoretic models are apt to be underspecified in the sense that they do not model what happens before the movement in the game begins. Thus, the most likely factors not included in a game theoretic model include the characteristics of the actors themselves and how the actors become involved in an interaction. However, the nature of the behavior and factors which are integral to the interaction, are typically specified in a game theoretic model. These factors prove to be extremely useful when specifying an empirical model.

Having illustrated the differences and similarities of each approach, it is important to realize that any methodology has its draw backs. In fact, no methodological tradition can explain all social behaviors. This is why a modeling dialogue is needed (Myerson 1992: 64). We need to recognize the strengths and benefits of modeling that can be derived from game-theory and vice versa.

Integrating the Formal Approaches

Although differences between game theoretic models and empirical models abound, as the previous discussion indicates, so do numerous commonalities. Again, it is our contention that much insight could be gained by integrating both formal approaches to political science. As stated previously, game theoretic models offer an approach to modeling which is very systematic and rigorous. Empirical specification, on the other hand, allows us to search for patterns to appraise how applicable our theories are in the context of the real world. Integration of the two formal approaches has been attempted, but only on a limited basis. In the discussion that follows we suggest that synthesizing the two approaches can take place in an inductive or deductive manner. Inductively, game theoretic models can be deduced which allow one to select among competing empirical models. Indeed, this is precisely what Morrow (1989) does with

[10] See for example Morton (1993: 382-392) or Palfrey (1991). See Green and Shapiro (1994) for a critique of this literature. Case studies also can be used to evaluate game theoretic models empirically (although not as rigorous).

a limited information model of crisis bargaining. Deductively, game theoretic models can be used as a tool in deriving a model's corresponding empirical specification. This permits the model to keep intact the analytic rigor characteristic of game theoretic models, while simultaneously incorporating the ability to test it empirically. We argue that the integration of both formal methods render tests of theories with models that are more analytically rigorous and empirically testable.

An Inductive Approach to Integration

An inductive approach consists of using game theoretic models to select among competing empirical models. Morrow (1989) accomplishes this with a formal model he develops to study crisis bargaining. In his study, Morrow develops a game theoretic model that analyzes how sequence and information compel actors to select certain actions in a crisis bargaining situation. He subsequently uses his game theoretic model to assess the viability of several empirical models of crisis bargaining. In this manner, Morrow demonstrates how game theoretic models can be used to develop testable propositions and to reexamine empirical studies. Morrow uses his formal model "to draw logical links between postulated underlying processes and empirical regularities. By demonstrating what empirical patterns should follow from an assumed process, the formal model serves as a tool to judge evidence" (1989: 964).

Specifically, Morrow uses his model to demonstrate how empirical models of crisis behavior are plagued by the joint problems of selection bias and misspecification. Both of these problems stem from the unobservable nature of beliefs. Morrow develops a sequential equilibrium of the model that characterizes the offers and acceptance of offers for both sides, for all moves within the game theoretic model. His model illustrates four basic cases which are a function of each side's initial beliefs, the status quo, and the costs of war for each side. They include (1) where one nation holds a military preponderance over another nation; (2) one nation always initiates a crisis; (3) neither side believes the other has an advantage; (4) and lastly, a "noncase" in which no crisis occurs.

From this game theoretic analysis, he proceeds to evaluate existing empirical models in an attempt to see which studies more appropriately characterize the nature of crises bargaining. Morrow's model systematically reveals that existing empirical models are characterized by misspecification problems and selection bias. In terms of selection bias, Morrow contends that crises situations are not typical dyadic relationships. However, most sample populations used to test empirical models consist solely of crisis situations. By failing not to control for "non-crisis" situations, any explanation regarding dyadic relationships is plagued by a selection bias problem. Moreover, he finds specification problems in empirical models as they fail to control for each sides' (actors') beliefs about the quality of military capabilities. These beliefs are important as they are crucial in determining what type of actions each side will pursue. This is partly because such beliefs are generally not observable, which complicates any attempt to include the omitted variable. (It is difficult to assess how each sides' real capabilities compare with each sides' perceived capabilities). In addition, the selection bias complicates the misspecification problem by creating a biased distribution of beliefs among the cases of crises (Morrow 1989: 958-964). The formal game theoretic model reveals that unless the selection effects are controlled, conclusions can not be drawn from the empirical models. Moreover, the formal game theoretic model illustrates how the crises observed in the empirical models are but a subset of all crisis situations, and merely account for a small proportion of dyadic relationships.

By viewing several empirical studies through the lens of his sequential game, Morrow is able to detect problems affecting existing empirical models. As a result, he is able to correct these problems so that new conclusions can be drawn. Indeed, this process highlights the real power of game theory, as it is used to generate new findings and understandings of our theories (Snidal 1985). Moreover, this emphasizes the benefits to be gained from integration. Namely, that competing empirical models can be evaluated for analytic correctness and rigor with a corresponding game theoretic analysis. The game analytic framework was able to show systematically what factors were important in explaining behavior in crises situations, and which factors were improperly being ignored. This type of formal integration can also prove to be useful for conducting encompassing tests. It has already been established that game theoretic analysis is able to rigorously specify relevant parameters in a model. Moreover, as the previous discussion suggests, it is also able to dismiss parameters which are largely irrelevant in empirical models. As such, if game theoretic analysis is used to assess irrelevant and relevant parameters in empirical models, in the same vein, it can be used to determine which models explain more. That is, game theoretic models can help identify and derive empirical models which explain the most and yet at the same time prove to be the most parsimonious.

A Deductive Approach to Integration

Alternatively, integration of the formal approaches can been done in a deductive manner. This process consists of using game theoretic models a priori as a theoretical devise to specify relevant parameters in a model. For example, Bueno de Mesquita and Lalman (1992) show how deductively integrating the two methods can be extremely useful. They illustrate how it is possible to take a theory about a particular phenomena, form a game theoretic model, and subsequently use the game theoretic model to shape their empirical test. They start with an international interaction game characterized by sequential decisions. The game features the actions of two rivals and the consequences of their strategic choices. From this basic game, Bueno de Mesquita and Lalman derive a set of propositions used to assess two competing theories of international relations: Realpolitik, which focuses on systemic constraints and the Domestic Politics variant, which incorporates both systemic and domestic political constraints in understanding strategic choice in international relations. These propositions are analyzed both formally and empirically through statistical analyses of 707 dyadic interactions. They conclude that the Domestic Politics model provides a better explanation of international relations. The analyses is deductive in that it begins with a general model, general propositions are derived, and finally propositions are empirically analyzed.[11]

Using the deductive approach, a game-theoretic model is developed and formalized to model some social interaction. In turn, assumptions (initial conditions) and equilibria for the game are identified. From this, a set of propositions are laid out which can be empirically tested. The empirical tests essentially test propositions that emerge from the analysis and identification of equilibria. What this suggests is that propositions can be extracted from game models which can be empirically tested in an effort to confirm or disconfirm the explanatory power of our theories. In this manner, a deductive approach provides a direct path of integration for game theory and empirical methods.

The deductive approach also provides an opportunity to test a variety of propositions derived from a set of formal models. As a variety of formal models provides a better understanding of the nature of specific social interactions, this type of comparative evaluation of propositions derived from a set of formal models contributes even more to our understanding. In fact, such a comparative evaluation of propositions may be one of the best ways to maintain theoretical analytical rigor and provide a more fully specified empirical model which can be statistically evaluated.

Conclusion

Game theoretic analysis has played a significant role in the field of Political Science. However, integrating game-theoretic and empirical methods still appears to be a novel idea. A move in this direction can only serve to advance the discipline as it can help our theories become more rigorous and systematic. If our aim as political scientists is to contribute to the discipline in this manner, then integrating these two analytic disciplines certainly is reaching towards that goal.

Both formal approaches have their unique ways for generating predictions. Myerson (1992) suggests that game theoretic predictions are theoretical output that empiricists need to guide their search for patterns in empirical data. We agree but take his suggestion a little further and suggest that game models can help specify actual parameters in an empirical model. Game theoretic models yield predictions which can be used as "theoretical output" for empiricists to guide them in specifying parameters and structures of empirical models. Both inductive and deductive forms of integration can help foster an effective dialogue between these two formal approaches. The combination of the two approaches can help develop systematic and rigorous models in political science.

[11] We should note that not all of Bueno de Mesquita and Lalman's analysis is purely deductive. (We also are aware that no research is ever purely deductive or inductive, but exhibit general tendencies in one direction or another). They also engage in more inductive analysis, examining such empirical puzzles as: why democracies do not fight one another, but are often engaged in conflict with other types of regimes?

References

Achen, Christopher H. 1992. "Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research," *Political Behavior*, v. 14.

Blackwell, David H. and Meyer A. Girshick. 1954. *Theory of Games and Statistical Decisions*. New York: Wiley.

Brehm, John and Scott Gates. 1992. "Policing Police Brutality," paper presented at the annual meetings of the Midwest Political Science Association, Chicago, IL.

Bueno de Mesquita, Bruce and David Lalman. 1992. *War and Reason*. New Haven: Yale University Press.

Dow, Jay and Melvin J. Hinich. 1992. "Estimating the Parameters of Spatial Models." Paper presented at the Ninth Political Methodology Conference, Cambridge, MA.

Gaddis, John L. 1987. *The Long Peace*. New York: Oxford University Press.

Gerber, Elisabeth and John E. Jackson. 1993. "Endogenous Preferences and the Study of Institutions," *American Political Science Review*, 87: 639-656.

Green, Donald and Ian Shapiro. 1994. *Pathologies of Rational Choice Theory: A Critique and Applications in Political Science*. New Haven: Yale University Press.

Johnson, James. 1992. "Is Talk Really Cheap? Prompting Conversation Between Critical Theory and Rational Choice," *American Political Science Review*, 87: 74-86.

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

Kreps, David M. 1990. *A Course in Microeconomic Theory*, Princeton: Princeton University Press.

Lalman, David, Joe Oppenheimer, and Piotr Swistak. 1993. "Formal and Rational Choice Theory: A Cumulative Science of Politics." In *Political Science the State of the Discipline II*, ed. Ada W. Finifter. Washington, D.C.: American Political Science Association.

Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining," *American Journal of Political Science*, 33: 941-972.

Morrow, James D. 1994. *Game Theory for Political Scientists*. Princeton: Princeton University Press.

Morton, Rebecca B. 1993. "Incomplete Information and Ideological Explanations of Platform Divergence," *American Political Science Review* 87: 382-392.

Myerson, Roger B. 1992. "On the Value of Game Theory in Social Sciences," *Rationality and Society* 4: 62-73.

Myerson, Roger B. 1991. *Game Theory. Analysis of Conflict*. Cambridge, MA: Harvard University Press.

Palfrey, Thomas R., editor. 1991. *Laboratory Research in Political Science*. Ann Arbor: University of Michigan Press.

Putnam, Hilary. 1985. "Reflexive Reflections," *Erkenntnis* 22: 143-153.

Quine, W.V.O. 1953. "Two Dogmas of Empiricism," *From a Logical Point of View*. Cambridge, MA: Harvard University Press, pp. 20-46.

Rasmusen, Eric. 1989. *Games and Information*. New York: Basil Blackwell.

Snidal, Duncan. 1985."The Game *Theory* of International Politics," *World Politics* 138: 25-57.

Wilson, Rick and Roberta Herzberg. 1994. "Laboratory Experiments as a Tool for Institutional Analysis and Design." Rice University mimeo.

Review of *Experimental Foundations of Political Science*, Donald R. Kinder and Thomas R. Palfrey, eds., University of Michigan Press, 1993.

Greg D. Adams
University of Iowa

Kinder and Palfrey make no secret of the intent of their edited volume, *Experimental Foundations of Political Science*; as they put it, their "aim throughout is the unabashed promotion of experimentation" (2). In pursuit of this goal, Kinder and Palfrey bring together fifteen articles, principally from the *APSR*, and assemble them into five chapters, each addressing a different area within political science. Diversity was clearly a major strategy in organizing the book. Among other topics, chapters address how groups are able to resolve collective action problems, the ways in which poorly informed voters can arrive at reasonable decisions, and how surveys can either clarify or mask the nature of public opinion.

The diversity of the articles in the volume has both plusses and minuses, however. On the one hand, the chances of coming across an experiment that addresses one's own particular interests are quite high (though significantly smaller for IR or comparative researchers). On the other hand, many of the articles will undoubtedly confront each reader's paradigmatic biases. Experiments in political science often stem from work in either psychology or economics, and the differences between the two schools can be stark. Thus, readers with a background in formal theory may find some of the psychology-based articles tedious and unfocused, while others trained in psychological techniques may find the economic-based articles to be dry or incomprehensible. So, although most readers will find some articles of interest, they will also be likely to find at least one or two that are disappointing. How the sum of the articles balances out in the reader's mind will depend on the individual's background and the strength of his or her biases. On average, people should expect to find about three-quarters of the articles worthwhile — not bad considering the variety of topics in the book.

There are some articles that almost every reader will appreciate though. Quattrone and Tversky's piece presents strong experimental evidence describing the nonlinearities in people's utility curves. Cover and Brumberg offer an innovative article in which they manipulate a (willing) member of Congress's mass mailings to determine the effectiveness of the franking privilege in Congress. And Levine and Plott recount a humorous "real world" incident in which they use agenda power to further their own interests in a flying club. Unfortunately, Kinder and Palfrey omit some novel experiments that I think could have improved the book. With the advent of powerful multi-media computers and high-speed computer networks, experiments are becoming easier to implement and are pushing the envelope of creative social science. Forsythe, Nelson, Newmann, and Wright (1992) offer a perfect example of this in an article in the *American Economic Review*, in which they describe the results of the Iowa Political Stock Market that was (and is still) run over the Internet. I would have preferred to see more of these novel approaches included in the volume.

In terms of fulfilling the book's agenda of promoting experiments, though, the strongest piece of the book is the preface. Kinder and Palfrey provide an excellent review of the experimental literature in political science, and they give a thoughtful summary of the virtues and vices of experimental methods. As the editors explain, experiments have the potential to answer questions that are not easily solved through other methods. They allow for greater control of causal factors, they are amenable to replication, and they create interdisciplinary bridges that otherwise might not exist. Experiments are also, however, more vulnerable to criticisms regarding their generalizability. Many experimenters routinely face charges that their "treatment" variables were artificially contrived, that behavior in the lab fails to mirror the real world, and that conclusions about a general population based on a small, atypical sample of subjects (often college sophomores) are at best tenuous. Kinder and Palfrey address these criticisms effectively, pointing out that these concerns are not entirely unique to experimental methods and that many experimental designs get around such problems. In fact, the differences between experimental and nonexperimental data are often overstated. In the words of econometrician Edward Leamer (1983), "The truly sharp distinction between inference from experimental and inference from nonexperimental data is that experimental inference sensibly admits a conventional horizon in a critical dimension, namely the choice of explanatory variables" (39). Given the complexity of social phenomena, such admissions are hardly unwelcome.

On the whole, *Experimental Foundations* achieves much of what its editors set out to accomplish. Like any edited volume, there are pieces that readers will appreciate more or less than others, and these will depend largely upon individual tastes. But almost every reader will come away from the book with a sense that experiments can produce powerful and interesting results for a variety of political questions. The book is an excellent beginning for those interested in expanding their methodological toolset to include experiments, and I would certainly recommend it as part of a course covering general research methods.

References

Forsythe, Robert, Forrest Nelson, George R. Newmann, and John R. Wright. 1992. "Anatomony of an Experimental Political Stock Market," *The American Economic Review*, 82: 1142–61.

Leamer, Edward E. 1983. "Let's Take the Con out of Econometrics," *The American Economic Review*, 73: 31–44.

Review of Arthur S. Goldberger, 1991. *A Course in Econometrics*, Cambridge Massachusetts; London, England: Harvard University Press, 405pp.

Gary King
Harvard University

Arthur Goldberger's *A Course in Econometrics* was reviewed by Clive Granger along with three other major econometrics books in a recent issue of the *Journal of Economic Literature* (March, 1994, Pp.115-122). Granger presented a long checklist of what he was looking for in each book, a list that incidentally would be envied by anyone seeking the latest in econometric buzzwords. It included the encompassing principle, AIC model selection, cointegration, Kalman filters, autoregressive conditional heteroskedastic models, chaos, Brownian motion, and all kinds of other models, methods, tests, and ideas. By Granger's measure, Goldberger's book fails miserably. According to Granger, Goldberger did not give even a single item "substantial coverage," and the vast majority were not covered at all.

Granger's checklist is useful, but if you are looking for a book to teach or learn from it misses the point. Learning the latest econometric model or method is sometimes useful. The problem is that recent studies indicate that a new econometric trick is invented every 23 minutes, and at least three of every four are not useful, bad advice, or not applicable to political data. One solution to the problem of choosing a text is to avoid giving students your favorite bag of tricks, but to instead teach them the fundamentals, an area where Goldberger's text is unequaled. This will enable students to ascertain whether each new wrinkle in the econometric literature has some basis in statistical theory or might be useful in application to our data. This approach also enables them to avoid fads, intelligently choose among methods, distinguish disciplinary conventions from basic statistical inference, and enable them to modify new methods to suit their purposes.

Goldberger's book is a formal presentation of lecture notes from his famous two semester class at the University of Wisconsin, Madison. His class is required of all entering economics graduate students and is also taken by a handful of frightened interlopers from other departments, including political science. *A Course in Econometrics* was published nearly a decade after I took the course, but its resemblance to my current lecture notes is remarkable.

Goldberger's approach might be called "practical theory." His book does not focus on data analysis, although it includes a few pages of tutorial on Gauss and an occasional example with real data. The book is also not a mathematics text, although it includes proofs of almost everything he asserts. Instead, his focus is on theory where it is useful to application. He insists on formulating theoretical models, but he does not take these models particularly seriously. For example, he teaches how to demonstrate whether an estimator is unbiased, but if he finds a biased estimator his goal becomes figuring out whether the feature of the joint probability distribution that you are estimating is interesting, even though it may not be the one you originally intended. Goldberger is also not particularly interested in causality, and instead asks readers to center their efforts on making inferences about the conditional expectation function, $E(Y|X)$. He is not even bothered if the conditional expectation is nonlinear and you run a linear regression. Instead, he shows how the regression can estimate the best linear approximation to the conditional expectation. In a manner more common among statisticians, even disturbance terms have no life of their own. They are defined merely as deviations from the conditional expectation, $y - E(Y|X)$.

A Course in Econometrics, and Goldberger's econometrics course, begins with probability theory. His preferred model of the world is a joint probability distribution, from which one can describe just about anything. He discusses expectations, variances, special distributions, and several ways variables can be related or unrelated in probability theory. He takes eleven chapters (as I recall, about eight weeks) before anything is added up and divided by n. He discusses least squares, instrumental variables, and maximum likelihood estimation as special cases of the "analogy principle," which provides an easy and perhaps the most obvious means of matching features of probability distributions with estimators. Most of the remaining twenty-three chapters introduce a variety of topics in linear regression analysis. He also introduces a few advanced topics in the closing chapters but much more briefly.

His presentation provides as much insight in as concise and organized a manner as possible, always stopping to show the methods he used to make his formal arguments. The book will not be easy for political science students but almost everything statistical they learn subsequently will be.

Review of *Predicting Politics: Essays in Empirical Public Choice* W. Mark Crain and Robert D. Tollison, eds. 1990, University of Michigan Press.[1]

Mitch Sanders
University of Rochester

The organizing principles of this compilation, according to the editors, are "the use of economics to explain politics and a belief that public choice theories should be tested." These are laudable goals, especially when pursued jointly, and in many ways the circumstances are favorable for this volume to fulfill them. Public choice theories exist to describe a wide range of institutional behavior, and there is a wealth of relevant information from state governments and Congress. The data are usually amenable to the methods of choice, OLS and discriminant analysis, so there are few difficult econometric problems. Unfortunately, though, the potential of the situation is largely unfulfilled, due to problems of theory, operationalization, or interpretation that bedevil most of the works in this volume.

In some essays, serious theoretical problems arise before any empirical models are attempted. "Legislative Majorities as Nonsalvageable Assets" models the size of the majority party in state legislatures without considering electoral results. "A Theory of Legislative Organization" assumes that 1) committees are random samples of the parent legislature, 2) the floor defers to committees, and 3) that there is no vote-trading. These are strong assumptions, requiring more extensive justification than the authors provide. "Bureaucratic Structure and Congressional Control" treats regulatory commissions as sets of multiple individuals, with no recognition of the effects of majority rule on the aggregation of preferences. "The Three Classes of Senate Seats" claims that legislative output will decrease when there is larger overlap between House and Senate constituencies, due to the increased difficulty of making deals across chambers. However, it is not clear why, in states with Senate races, deals cannot be made that benefit a majority of constituents in all House districts (and hence a majority at the state level), thus enhancing cooperation.

Other essays in this volume contain provocative ideas, but provide data that bear only a tenuous relationship to the relevant theoretical concepts. The authors of "Pork-Barrel Paradox" point out that political parties might limit spending on pork-barrel projects in "politically volatile" districts, where the relatively high probability that the seat will

[1] This compilation consists of nineteen essays by various permutations of twelve authors, with at least one of the editors involved in all but one of the essays. Eight selections were published previously.

change parties reduces the present value of federal spending. One interesting corollary of this idea is that term limits, which increase the probability of seat turnover, should also reduce the level of spending by reducing its present value. But the empirical model estimates the relationship between the electoral vulnerability of a district and the success rate of bills sponsored by that district's incumbent, and does not include any direct indicator of federal spending at the district level. "Behind the Veil: The Political Economy of Constitutional Change" argues that in the absence of an independent judiciary to enhance the durability of legislative contracts, interest groups will tend to pursue wealth transfers via amendments to the state constitution rather than through ordinary legislation. But the dependent variable is the change in the length of the state's constitution (in words), without consideration of how many of those words are parts of sentences or paragraphs dealing with wealth transfer. A similar idea operates in "The Executive Branch in an Interest-Group Theory of Government," where executive vetoes (to prevent the legislature from reneging on previously made agreements) are more likely in the presence of large majorities in the legislature and in the absence of an independent judiciary. It is interesting to note that this result does not depend on which party has a majority in the legislature. But the dependent variable in this model is simply the number of vetoes, without regard to the content of the bills vetoed. These essays have interesting theoretical results, but the empirical tests fall short by failing to account for the fact that not all bills, words or vetoes are created equally.

In other cases, theory and testing proceed effectively, but the interpretation of results is problematic. In "Expressive versus Economic Voting" the authors test these two alternative conceptions of the voting act using a model of state welfare expenditures. This is more plausible than it sounds, but still subject to the usual caveats about using aggregate models to make inferences about individual behavior. "An Economic Theory of Redistricting" indicates that judicial control of reapportionment is less likely to result in a majority party losing its majority status than legislative control of reapportionment, because courts are interested in the durability of legislation while individual members of the legislature are at least somewhat willing to trade majority status for enhanced electoral security. The authors claim support for their theory as applied to State Houses because the coefficient relating "Apportionment Agent" to "Majority Size" is significant at the .30 level. The authors do admit that this is "weakly significant," but the mind (at least my mind) boggles at the thought of the consequences of such a lax standard of inference.

The best of these essays manage to traverse the minefield of theory, evidence, and inference, and make significant contributions to our understanding of political institutions. "Democracy and the Marketplace" uses data from the states in 1915 to show that secret ballot requirements served to increase per capita state government expenditures, as the wealthy were unable to purchase "temporary majorities" to avoid wealth redistribution. "Laissez-Faire in Campaign Finance" argues that restrictions on campaign finance at the state level will decrease the value of passing additional laws by reducing the pool of contributors, but will increase the level of fiscal transfers as exporting costs to other districts becomes more attractive. In both of these cases solid theories are tested with appropriate data, producing worthwhile results.

Overall, this volume contains some interesting theoretical insights, and empirical applications that remind us of just how challenging it is to find and use appropriate evidence. In the virtues and flaws of these essays we can see the promise, and the difficulty, of developing and testing theories of economic behavior in political institutions.

1995 Preliminary American Political Science Association Methods Program Listings

John Williams
Indiana University

Program Division: Political Methodology

DYNAMIC MODELS AND RATIONAL CHOICE

Panel Chair: Jim Granato
Department of Political Science
Michigan State University
303 S. Kedzie Hall
East Lansing, Michigan 48824-1032
phone: 517-353-7886

Papers:

"Toward a Computable Political-Economic Equilibrium Model"

John R. Freeman
Department of Political Science
University of Minnesota
Minneapolis, Minnesota 55455
phone: 612-624-6018

"A Rational Choice Defense of Sims-VAR Economic Policy Analysis"

Robert Grafstein
Department of Political Science

University of Georgia
Baldwin Hall
Athens, Georgia 30602-1615
phone: 706-542-2057

"Economic Conditions and Endogenous Elections"

Harvey Palmer
Department of Economics
George Mason University
Fairfax, Virginia 22030-4444
phone: 703-993-1124

Guy Whitten
Department of Political Science
Texas A&M University
College Station, Texas 77843
phone: 409-845-6782

Discussant 1. Jim Granato
Department of Political Science
Michigan State University
303 S. Kedzie Hall
East Lansing, Michigan 48824-1032
phone: 517-353-7886

Discussant 2: Greg Adams
Department of Political Science
University of Iowa
Iowa City, Iowa 52242
phone: 319-335-2070

MEASUREMENT, PUBLIC OPINION, AND PUBLIC POLICY

Chair: Charles Franklin
Department of Political Science
University of Wisconsin, Madison
Madison, Wisconsin 53706
phone: 608-263-2414

Papers:

"How Social Desirability Biases and Administrative Procedures Can Affect the Validity of Administrative Reports and Self-Reports of AFDC Behavior"

Henry Brady & Samantha Luks
Department of Political Science
University of California, Berkeley
Berkeley, California 94720
phone: 510-486-1953

"Racist-Nationalist, Postmaterialist and Economic Motivations for Social Movements in France, the United Kingdom, and Belgium: A Longitudinal Structural Relations Analysis: 1984-1993"

Jasjeet S. Sekhon & Walter R Mebane Jr.
Department of Government
Cornell University
Ithaca, New York 14853
phone: 607-255-3868

"Conflicted Responses to Balanced Survey Questions: Another Look at Non-Response"

Jill Glather
Department of Political Science
State University of New York, Stony Brook
Stony Brook, New York 11794-4392
phone: 516-632-7655

Discussant 1: Charles Franklin
Department of Political Science
University of Wisconsin, Madison
Madison, Wisconsin 53706
phone: 608-263-2414

TOPICS IN TIME SERIES ANALYSIS

Chair: Renée M. Smith
Department of Political Science
University of Rochester
Rochester, New York 14627
phone: 716-275-3225

Papers:

"Time Series Analysis of Event Counts: Point Processes and Statistical Models"

Chi Huang
Department of Political Science
University of Kentucky
1615 Patterson Office Tower
Lexington, Kentucky 40506-0027
phone: 606-257-3136

Todd G. Shields
Department of Political Science
University of Arkansas
Fayetteville, Arkansas 72701
phone: 501-575-2642

"Testing Causality in the Presence of Deterministic and Stochastic Trends"

Paul Kellstedt
Department of Political Science
University of Minnesota
Minneapolis, Minnesota 55455
phone: 612-624-4144

"Coping with Measurement Errors Over Time: The Estimation of a Dynamic Shock-Error Model"

Gregory E. McAvoy
Department of Political Science
Duke University
Box 90204
Durham, North Carolina 27708-0204
phone: 919-660-4320

Discussant 1. Renée M. Smith
Department of Political Science
University of Rochester
Rochester, New York 14627
phone: 716-275-3225

Discussant 2. Janet Box-Steffensmeier
Department of Political Science
Ohio State University
Columbus, Ohio 43210
phone: 614-292-2880

ISSUES IN SPECIFICATION

Chair: Chris Achen
Department of Political Science
University of Michigan
5601 Haven Hall
Ann Arbor, Michigan 48109
phone: 313-764-6312

Papers:

"Specification Uncertainty and Model Mixing"

Larry Bartels
Department of Politics
Princeton University
Princeton, New Jersey 08544-1012
phone: 609-258-4794

"Bayesian Tools for Social Scientists: Gibbs Sampling, Data Augmentation, and EM"

Simon Jackman
Department of Political Science
University of Chicago
5828 South University Avenue
Chicago, Illinois 60637
phone: 312-702-8075

"The Ecological Veracity"

Gary King
Department of Government
Harvard University
Cambridge, Massachusetts 02138
phone: 617-495-2027

Discussant 1: John R. Freeman
Department of Political Science
University of Minnesota
Minneapolis, Minnesota 55455
phone: 612-624-6018

MODELING CHOICE

Chair: John Aldrich
Department of Political Science
Duke University
Box 90204
Durham, North Carolina 27708-0204
phone: 919-660-4346

Papers:

"Specifications of Models of Individual Choice: Theoretical and Empirical Considerations"

R. Michael Alvarez
Division of Social Sciences
California Institute of Technology
Pasadena, California 91125
phone: 818-395-4273

Jonathan Nagler
Department of Political Science
University of California, Riverside
Riverside, California 92521
phone: 714-787-5502

"Quantal Response Equilibria in Extensive Form Games"

Richard D. McKelvey & Thomas Palfrey
Division of Social Sciences
California Institute of Technology
Pasadena, California 91125

phone: 818-405-9841

"Centripetal Truncation in Models of Electoral Choice"

Mitch Sanders
Department of Political Science
University of Rochester
Rochester, New York 14627
phone: 716-275-3190

Discussant 1: John Aldrich
Department of Political Science
Duke University
Box 90204
Durham, North Carolina 27708-0204
phone: 919-660-4346

METHODOLOGICAL ISSUES IN COMPARATIVE AND INTERNATIONAL POLITICS

Chair: Mark G. Lichbach
Department of Political Science
University of Colorado at Boulder
Boulder, Colorado 80309
phone: 303-492-4586

Papers:

"Traveling Methods: Use and Potential Abuse of Roll Call Voting Analysis"

Moshe Haspel
Department of Political Science
Emory University
Atlanta, Georgia 30322
phone: 404-727-4586

"Recovering Events from Events Data: Modeling the Bias in Reuters' Coverage of Routine-Political and Direct Action"

Karen Rothkin
Department of Political Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
phone: 617-891-9956

Doug Bond
Department of Government
Harvard University
Cambridge, Massachusetts 02138
phone: 617-495-5580

"Rationality and Comparative Method"

Junko Kato
Department of Social Sciences
University of Tokyo
3-8-1 Komaba Meguro-ku
Tokyo, Japan 153
phone: 011-81-3-5454-6460

Discussant 1: Mark G. Lichbach
Department of Political Science
University of Colorado at Boulder
Boulder, Colorado 80309
phone: 303-492-4586

Discussant 2. Brian K. Collins
Department of Political Science
Indiana University
Bloomington, IN 47405
phone: 812-855-9690

ECOLOGICAL INFERENCE

Chair: John Sprague
Department of Political Science
One Brookings Drive
Washington University
St. Louis, Missouri 63130-4899
phone: 314-935-5810

Papers:

"Statistical Models for Inferring Individual Voter Instability from Aggregate Electoral Volatility"

Chris Achen
Department of Political Science
University of Michigan
5601 Haven Hall
Ann Arbor, Michigan 48109
phone: 313-764-6312

"Aggregate Data and the Voting Rights Act"

Wendy K. Tam
Department of Political Science
University of California, Berkeley
Berkeley, California 94720
phone: 510-642-1474

"Aggregate and Individual-Level Interaction Effects in Ecological Regression"

Brad Palmquist
Department of Government
Harvard University
Cambridge, Massachusetts 02138
phone: 617-495-9564

Discussant 1: John Sprague
Department of Political Science
One Brookings Drive
Washington University
St. Louis, Missouri 63130-4899
phone: 314-935-5810

Discussant 2: Christopher Kenny
Department of Political Science
Louisiana State University
Baton Rouge, Louisiana 70803-5433
phone: 504-388-2439

INNOVATIVE APPLICATIONS IN AMERICAN POLITICS

Chair: R. Michael Alvarez
Division of Social Sciences
California Institute of Technology
Pasadena, California 91125
phone: 818-395-4273

Papers:

"Nested Multinomial Logit Panel Models for Party Identification, Vote Choices and Local Economic Conditions: 1972-76 and 1980"

Suzanne M. Smith
Department of Government
Cornell University
Ithaca, New York 14853
phone: 607-256-4286

"Social Network Methods in the Study of Political Influence"

David Lazer
(University of Michigan)
182 Hudson Avenue
Red Bank, New Jersey 07701
phone: 908-758-6571

"War Chests and Challenger Quality"

Jay Goodliffe
Department of Political Science
University of Rochester
Rochester, New York 14627
phone: 716-275-5421

Discussant 1: David A. Scocca
Department of Political Science
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599-3265
phone: 919-929-8528

Discussant 2: Bradford S. Jones
Department of Political Science
University of Arizona
Tucson, Arizona 85721
phone: 602-621-3315

Political Analysis News

Under the editorship of John R. Freeman, *Political Analysis* is taking submissions for future issues. Send your submissions to John R. Freeman, Editor, *Political Analysis*, Department of Political Science, 1414 Social Sciences Building, University of Minnesota, Minneapolis, MN 55455.

Political Analysis is the primary vehicle for publication of research in political methodology, and as such, deserves your continued support. It is *the* journal of the Political Methodology Section of the American Political Science Association. With Volume 5 available early this summer (June 1995), it is time to make sure that your university library has a standing order for the journal of **YOUR** organized APSA section.

Also, the advertisement in this issue of *The Political Methodologist* has ordering information so that you can add *Political Analysis* to your personal library at special discount prices for members of the Section. The Editors would also like to point out that Malcolm Litchfield, Editor for Political Science at The University of Michigan Press, has recently clarified the procedures for individual standing orders for future issues of *Political Analysis*, a subject which has been discussed at recent gatherings of political methodologists. You can contact Linda Rowley (313-936-0396) to check on your standing order or to establish one.