The Political Methodologist

Newsletter of the Political Methodology Section,
American Political Science Association
Volume 2, Number 2, Fall 1989.

Contents

Notes from the Editor 1

Review of Converse and Pierce, *Political Representation in France, Christopher H. Achen* 2

Systematic Sampling and Temporal Aggregation: An Annotated Bibliography, *John R. Freeman* 4

Graphic Displays in Political Science, *Charles H. Franklin* 7

The 1989 Political Methodology Meetings, *Larry Bartels* 9

"Creating" Good Problem Sets, *John E. Jackson* 10

Choosing a Computer Package for Political Methodology Courses, *Nathaniel Beck* 12

An Undergraduate Syllabus, *George E. Marcus* 14

An Annotated Bibliography on Artificial Intelligence, *Philip A. Schrodt* 17

Job in Political Methodology 22

Notes From the Editor

We inaugurated *The Political Methodologist* last Winter to facilitate communication in the emerging field of political methodology. Judging by the hundreds of letters I received about the last issue from all over the world, communication has been established. Interest in the field also appears to extend well beyond the boundaries of the United States and American political science. It remains a mystery how the first issue found its way to numerous scholars in so many countries not on our mailing list.

Christopher Achen's thoughtful review of Philip Converse and Roy Pierce's monumental book, *Representation in France*, leads off this issue. John Freeman provides an extremely useful annotated bibliography on aggregation problems in time series models. Charles Franklin offers valuable advice about statistical graphics, and Larry Bartels reports on the sixth annual meeting of the Political Methodology conference.

Since the success of political methodology often depends on providing scholars in other fields of the discipline new ways to think about and analyze substantive problems, we have a special responsibility to be good teachers. In this issue, John Jackson describes the perfect solution to finding problem sets for class, and Nathaniel Beck gives his sage advice on the merits of various computer programs. George Marcus provides a syllabus for an undergraduate social science statistics class, and Philip Schrodt offers an annotated bibliography and a brief course syllabus on artificial intelligence. An advertisement for a job in political methodology appears last.

In the next issue of the newsletter, look for important contributions by Henry Brady, on numerical optimization, Nancy Burns, on a survey of political methodology requirements in graduate programs, and many other features. One special item will be an article by Harold F. Gosnell about the genesis of his classic sample survey study *Non-Voting* (1924), experimental work *Getting Out the Vote* (1927), and statistical analyses of voting behavior.

I encourage contributions of any kind related to political methodology, defined broadly. New items might include brief research reports, requests for data, methodological critiques of published articles, or unusual methodological problems. Please send contributions to me at the Department of Government, Harvard University, Littauer Center, Cambridge, MA 02138 (BITnet: gmk@harvunxw; Internet: gmk@wjh12.Harvard.Edu; FAX: 617-495-0438). We prefer submissions in TEX or LATEX formats on MS-DOS diskettes, but most other electronic formats will do. Subscriptions to the *The Political Methodologist* are free to members of APSA's Political Methodology Section and $15.00/year to others. *Gary King, Harvard University* TPM

Review of Converse and Pierce, *Political Representation in France,*

Christopher H. Achen, Department of Political Science and National Opinion Research Center, University of Chicago

Philip E. Converse and Roy Pierce. 1986. *Political Representation in France.* Cambridge, Massachusetts: The Belknap Press of Harvard University Press. 996 pages.

From its inception, the "Michigan model" of the voter's choice, particularly as set out in *The American Voter*, has been criticized for over-generalizing from America in the Eisenhower years. Other periods and countries, it is said, are less quiescent and more ideological. Above all, the key Michigan concept of "party identification" was not expected to suit sophisticated Europeans any better than other small-town Midwestern ideas.

The present volume is a massive and masterly defense of Michigan against these critics. Manifestly, it is an over-time study of French political opinions and behavior, at both mass and elite levels, during the late 1960's. But its informing spirit is a desire to replicate the key findings of *The American Voter, Elections and the Political Order*, and "The Nature of Belief Systems in Mass Publics" in an apparently unsuitable nation. If the critics are wrong about France, the authors seem to say, then they are probably wrong quite generally. If the Michigan model can be made to work in the *locus classicus* of ideological rigor and fragile parties, then it will surely apply widely across the democracies.

To this end, the reader is offered nearly eight hundred pages of text and more than two hundred pages of appendices. The prose sometimes goes on a bit, like an abstracted lecturer slowly circling the point on a warm spring afternoon. But the full case is made—and in the main, skillfully and powerfully. Michigan may be aging and less alluring these days than her competitors from the rational choice school, but one doesn't learn everything in school. And as this volume makes clear, she isn't about to go quietly.

No short review can do justice to the richness of the book, but a selection of propositions from it may convey its style and intellectual *tendance*. For example, Converse and Pierce note that the fractionalization of the French party system, the impermanence of its parties, and the traditional reluctance to discuss politics within the family combine to make acquisition and inter-generational transmission of party ID difficult. Yet they argue that many voters nevertheless acquire one, and these French citizens behave like British and American identifiers:

> Thus the two truly essential properties of the party identification phenomenon—individual stability over time and a marked retardation of change in party voting when an identification is present—turn up unquestionably in the French case. (p. 96)

An obvious competing explanation of French voting is ideological. Perhaps in the land where the left-right distinction originated, ideological placement becomes more powerful than party, or perhaps even eliminates it altogether as an explanation? Using familiar reasoning and methods, Converse and Pierce reject this counter-argument:

> Although more French respondents locate themselves in left-right terms than in terms of party loyalties, we do not have to go very far beneath the surface to learn that many of these self-placements are of questionable pedigree, since the individuals involved often have but a limited understanding of what the labels 'left' and 'right' mean politically, and many seem to choose the exact midpoint of the left-right continuum as a means of avoiding a more integral commitment. (p. 149)

Converse and Pierce make the case that the usual left-right dimension, interpreted as an economic or class-based cleavage, has limited explanatory power in the voting booth. Religion works better, they say, and in the period under study, perhaps also Gaullism. Going a bit beyond their argument, one might argue that in postwar France, Gaullism and anti-Gaullism served as a limited surrogate for the persistent two-party systems of the Anglo-American democracies. Thus voters behaved quite reasonably in orienting themselves toward this aspect of their politics rather than toward the often transitory political parties. The French voter, too, is no fool.

The book also includes some discussion of that old devil, attitude instability. Converse and Pierce argue, as expected, that differences in over-time stability of survey responses reflect genuine differences in understanding. A nice feature of their dataset is that mass and elites (candidates for deputy) were asked identically-worded policy questions in some cases, and some elites were interviewed more than once. Thus truly comparable over-time stabilities can be computed for both citizen and representative.

Converse and Pierce make much of the fact that the elite is more stable than the mass. They seem to believe that this is a major blow to their critics (e.g., this reviewer). But no prominent researcher, nor any one who has ever walked a precinct, disputes the existence of elite-mass differences. Converse and Pierce quote no one who does. Indeed, my own work discovered some real improvements in stability among the top-most sophisticated even within a mass sample. Elite-mass differences simply are not much in dispute.

The key question instead is whether survey questions, by their nature, have a good deal of vagueness in them, so that average citizens with a reasonable layman's grasp of politics will be made to look ignorant under the Converse-Pierce analysis. In my (perhaps all too predictable) judgment, nothing in this book really supports their view that the bulk of citizens constitute an ill-informed "marais," and some of the evidence points the other way. For example, even the highly politically involved deputies have average squared over-time correlations of just 0.61, indicating that nearly 40 percent of the over-time variance in the policy questions is noise (p. 251). If one corrected for that in the mass sample and in addition used measures which, unlike correlations, do not penalize citizens for being less extreme on average than deputies, the bulk of the mass sample would look, perhaps not brilliant, but at least creditable.

The book also contains a long, sophisticated discussion of the French two-round electoral system and voter behavior under it. Much of the evidence goes in the expected direction: many voters support their party, their religion, or their ideology at the first ballot, and then shift to the best available alternative at the second ballot. But Converse and Pierce argue that "best available" need not mean "closest on the left-right dimension." They find, for example, that hard-core Socialists without a candidate of their own at round two may often spoil their ballots rather than vote for the Communists, whom they resemble ideologically (at least in a unidimensional world).

The statistical analysis in the book is usually quite good, and often truly excellent. This section, however, is marred by a few idiosyncrasies, such as a tendency to percentage cross-tabulations in unexpected ways. We learn, for example, how far the average ballot-spoiler was from her best available choice, but not the more causally relevant fact of how many voters at a given ideological distance spoiled their ballot. Similarly, heavy emphasis is sometimes placed on the interpretation of cells in cross-tabulations which have fewer than twenty observations. On pages 354-55, for example, rational-choice arguments by Rosenthal and Sen are attacked by comparing certain mean distances on an ideological scale. These numbers differ from each other by fewer than five points, and no 95 percent confidence intervals are given. However, the intervals are computable from other information in the text: one of them covers almost twenty points, and a second stretches nearly to thirty. Here and there throughout the book, one wishes more confidence bands were constructed and more significance tests done.

Particularly in the chapters on voting, the style of explanation is social-psychological and micro-level, perhaps even a bit old-fashioned, reflecting the period in which the fieldwork was done. Party positions and competitive strategies are taken as exogenous, and their variation over time rarely figures in the analysis. We get little sense of how the election campaigns happened, for instance. Party ID is generally treated as fixed; retrospective evaluations play almost no role. The state of the economy is pretty much ignored.

Explanation occurs at the level of single voters or candidates, and "attitudes" are given pride of place as causal factors. Much effort is expended to show that rational choice explanations based on perceived ideological distances work less well. For example, in a discussion of second-ballot choices by first-round centrist voters, we read (p. 375):

> The Gaullist supporters were distinctly favorable to Gaullism, held a favorable opinion of the Gaullist party, and were hostile toward the leftist party in their district. The leftist supporters, on the contrary, were anti-Gaullist, negative toward the Gaullist party, and moderately favorable toward the leftist party they voted for.

Or again in a treatment of second-ballot endorsements by certain socialist candidates who dropped out after the first round (p. 399):

> The Federation candidates who supported the Communists had, on average, a favorable attitude toward the Communist party, and those who did not help the Communists were distinctly anti-Communist. It is as simple as that.

Quite apart from their neglect of systemic factors, explanations of this kind can be criticized as tautological. If "attitude" were replaced by "utility" in the above quotations, so that actors were said to make choices because they had a utility for doing so, the explanation would obviously lack bite. Converse and Pierce's claims, and traditional social-psychological approaches to voting generally, have content only if attitudes have independent causal standing and are not simply another way of reporting dispositions to act in certain ways—not just revealed preferences, as the economists would have it. Converse and Pierce recognize this point, and they are firm (p. 681):

> For attitudes, values or other preferences, the proof of the pudding is often taken to be the degree of fit between verbal responses defining such preferences and other evidence as to actual choice of behavior. The assumption that attitudes are unreal except as they have some outcome in motor behavior is rather glib, and one to which we do not entirely subscribe, since there are any number of reasons, in addition to "unreal" attitudes, why relevant behaviors might not be carried out.

The book concludes with a detailed and thorough discussion of representation in France—its causal logic and systemic consequences. The dataset Converse and Pierce have assembled for this purpose is, quite simply, fabulous. The analysis is guided primarily by the logic of Warren Miller and Donald Stokes' work, and the same representation diamonds are constructed (linking constituency views, the deputy's perception of them, his own views, and his roll calls). A sample of the findings: deputies from safe seats are more representative than those from marginal districts, foreign policy views are represented no less well than other issues, a deputy's party is the best predictor of roll call votes, and the deputies' supporters in their districts are better represented than their constituencies as a whole.

The reader is treated as well to a sophisticated discussion of modes of representation and the strengths and weaknesses of various empirical indices. After trying a variety of measures (including some suggested by this reviewer), they put heaviest emphasis on the simple correlation between representatives and the mean opinion of the represented (corrected for error due to small constituency samples). Their conclusions depend heavily on this approach.

Such a procedure may often work tolerably well, but the correlation has weird properties in this context due to its dependence on the cross-district heterogeneity of opinion. For example, if the United States annexed a large number of white South Africans, and if they in turn imported their current representatives to become members of Congress, the representativeness of the Congress on racial matters would surely not be improved: every American and South African would have precisely the same representative as before, and every district's mean would be farther from the mean of the new combined legislature than it was from the mean of its old national legislature. Yet "representativeness" as measured by the correlation coefficient would take a big positive jump—the more variance in opinion across districts, the higher the correlation.

Converse and Pierce defend themselves on the point by arguing, entirely correctly, that proximity of views to one's own representative does not exhaust the meaning of representativeness. The rate of change in deputy positions as a function of constituency opinions is a key aspect of representation as well. But that is an argument for using the slope, not the correlation, particularly since the slope has no difficulties with changes in variance among independent variables.

This is no technical quibble. Unfortunately, cross-district variances fluctuate wildly from one issue to the next and from one definition of the represented to another. Correlations will move with them, even if the underlying causal process is unchanging. How many of the book's findings on representation depend on the odd properties of the correlation coefficient remains an open question.

As this review demonstrates, the temptation to argue with this volume is strong. For some devotees of the New Age research techniques, who will wish the book more *au courant*, the temptation to ignore it will be even stronger. Yet one suspects that Converse and Pierce will hold up well over time against the trendy critics. For *Representation in France* is the old time religion. The data are honored, the fancy techniques and theories are questioned, the political conventional wisdom is shattered. No study of a nation's representation system, so comprehensive and monumental, has ever been produced before, certainly not in a single volume. The sheer magnitude of the enterprise is astonishing, and the volume of powerful evidence and argumentation will set the agenda in behavioral studies of French politics for many years to come. TPM

Systematic Sampling and Temporal Aggregation: An Annotated Bibliography, *John R. Freeman, University of Minnesota*

[EDITOR'S NOTE: This annotated bibliography is part of John Freeman's scholarly work on the subject. For a very interesting introduction to these and other time series problems, I recommend his article, "Systematic Sampling and Temporal Aggregation," in Volume 1 of Political Analysis.*]*

Brewer, K.R.W. (1973) "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models," *Journal of Econometrics* 1, pps. 133-154.

Demonstrates the effects that systematic sampling and temporal aggregation have on the orders of these two kinds of models, for instance, how systematic sampling at interval k transforms a ARMA(p,q) process into a ARMA(p,r) process where r is the integer part of $[p+(q-p)/k]$ and temporal aggregation changes the former model into an ARIMA(p,r) process where r is the integer part of $[k\{(p+1)(k-1)+q\}]$. Contains a few errors see Weiss (1981) and it is a bit out of date (see Weiss, 1984). Useful nonetheless in explaining the mathematical reasoning on which the derivations of the effects of the two measurement practices are based.

Geweke, J. (1978) "Temporal Aggregation in the Multiple Regression Model," *Econometrica* 46(3), May, pps. 643-661.

Extends Sim's (1971) results on discrete approximations of continuous time bivariate relationships to the multivariate case. Shows that for this more general model, the macro coefficients will be a complex combination of the underlying (continuous) coefficients, or that the aggregated coefficients will be "contaminated" by aggregation. Suggests a procedure (for determining if one's estimates are plagued by this problem for cases in which one has data at several levels of aggregation). Ends with illustration, distributed lag model for the determinants of wholesale prices. Analysis will be difficult to follow for those not conversant in spectral methods. However, the article contains useful summaries of consequences of temporal aggregation, and Geweke's procedure for gauging the effects of temporal aggregation appears to be of some practical value.

Granger, C.W.F. (1988), "Aggregation of Time Series Variables—A Survey." Discussion Paper 1. Institute for Empirical Macroeconomics. Minneapolis, MN: Federal Reserve Bank of Minneapolis.

Reviews the results on the effects of small and large scale (contemporaneous) aggregation and of temporal aggregation, including his own new work on the concept of "common factors." Discusses implications of the three practices for causality testing, forecasting, and cointegration. Not as thorough a review as works like Weiss (1984), although more up to date. Contains an interesting discussion of the implications of (large-scale) aggregation of time series corresponding to decisions of millions of individuals and firms, e.g., aggregation of millions of consumption decisions. Actually more useful as review of the consequences of contemporaneous than temporal aggregation of time series.

Hawawini, G. and A. Vora (1983) "Temporal Aggregation and the Strength of the Association Between Securities' Risk and Return," *Economic Letters* 11, pps. 269-278.

Studies the effect of temporal aggregation on the R statistic for a financial model, more specifically, for the correlation between securities' average returns and the systematic risk as specified by the capital asset pricing model. Shows that this measure of fit is affected by the practice of averaging rates of return over various number of days. Includes illustration for sample of 1115 U.S. securities in the period 1962-1976.

Hawawini, G. (1978) "A Note on Temporal Aggregation and Serial Correlation," *Economic Letters* 1, pps. 237-242.

Shows how the first order serial correlation coefficient from a random walk model of stock prices is a complex function of the underlying (true) serial correlation coefficients, and how the first order coefficient goes to zero as one aggregates beyond the highest order serial correlation in the underlying data. Implies that researchers' repeated finding of random walks for stock prices could be an artifact of the aggregation of time series data across the true interval of the pricing process. Of general interest with respect to the detection and interpretation of autocorrelated errors; of special interest to users of weakly specified models and the proponents of rational expectations theories of politics (theories which purport to explain the repeated finding of random walks in international relations and other fields of political science).

Lutkepohl, H. (1984a) "Linear Aggregation of Vector Autoregressive Moving Average Processes," *Economic Letters* 14, pps. 345-350.

Presents general multivariate mathematical formulation of the practices of systematic sampling and contemporaneous and temporal aggregation on vector ARMA processes. Shows that the practices amount to linear transformations of the associated *macro* vector ARMA processes. Proves a theorem about upper bounds of the orders of the general models that are obtained from the three measurement practices. Primarily of theoretical interest.

Moriguchi, C. (1970) "Aggregation Over Time in Macroeconomic Relations," *International Economic Review* 11(3), pps. 427-439.

Precursor, in some ways, to Tiao and Wei (1976) and Wei (1978). Studies effects of temporal aggregation on simple, bivariate finite and infinite distributed lag models under certain assumptions about the character of the right-hand-side variable, x, more specifically, under the assumptions that x is fixed, x is a linear or exponential trend, and x has seasonal and (or) "irregular" components. Derives consequences for calculation of average time lag of response of y to x (weighted mean of relevant lags), biasedness, and efficiency. Actually develops formula for efficiency loss from temporal aggregation for a simple bivariate model with no lags or right-hand-side endogenous variables. A bit more difficult than Zellner and Montemarquette (1971) because, among other things, Moriguchi uses fractional lags for his aggregated models. But, generally speaking, a good introduction to the literature and an introduction with some practical value insofar as the characterization of the x variable corresponds to cases that political scientists actually encounter in their research.

Sims, C. (1971) "Discrete Approximations to Continuous Time Distributed Lags in Econometrics," *Econometrica* 39(3), pps. 545-563.

Examines conditions under which a bivariate, discrete time distributed lag model is a good approximation of the underlying continuous (bivariate) model, or when discrete point samples on the two variables can be effectively used to identify the structure of the true process. Shows that a good approximation can be obtained if the right-hand side variable is sufficiently "smooth, that is, if the right-hand side variable has certain frequency domain or spectral properties. Also considers aggregated version of the Koyck exponential and polynomial lag models. A very difficult article if one is not conversant in spectral methods. In addition, implications difficult to gauge since, unlike economists, we have few studies that survey the "typical spectral shapes" of political variables.

Stram, D. and W. Wei (1981), "Change of Model Form Under Temporal Aggregation in the ARIMA Process," Proceedings of the Business and Economics Section, American Statistical Association, pps. 313-317.

Derives exact formulae for orders of temporally aggregated models (not just upper bounds or maximum orders) under certain conditions pertaining to the identity and multiplicity of the roots of the underlying (true) ARIMA process. Also shows that for IMA models, many different, aggregate models can be obtained, that is, unlike the AR case, there is no simple relation between the roots of the MA polynomials in the true and aggregated models. That the order of the MA portion of the model for the aggregates can be less that the derived maximum order is not a consequence of the cancelling of roots as in the case of the AR portion of the model. Analysis is somewhat difficult to follow. But implications of this analysis are quite important for those who want to draw inferences about natures of political processes from the structures of fitted time series models, e.g., the proponents of rational expectations theories of politics.

Tiao, G. and W. Wei (1976), "Effect of Temporal Aggregation on the Dynamic Relationship of Two Time Series Variables," *Biometrika* 63(3), pps. 513-523.

One of the central works on the subject. Derives formulae for the Direct Aggregate and the conditional Aggregate of a one sided, finite distributed lag model, and shows that temporal aggregation, under most conditions, will transform the one-way causal relationship into a two-way or feedforward relationship. In this context demonstrates that the estimation of the usual (misspecified) temporally aggregated equation will ignore correlation between the (aggregated) right-hand-side variable and the (aggregate) error term, and, in turn, produce inconsistent estimates of the true or underlying effects of the right-hand-side variable on the dependent variable. Includes a mathematical example that shows information loss that occurs form temporal aggregation and that also demonstrates how the one-way causal relationship is transformed into a feedback relationship. Difficult reading but extremely important insofar as it contains one of the most important results about the way temporal aggregation undermines causal inference.

Wei, W. (1982), "Comment: The Effects of Systematic Sampling and Temporal Aggregation on Causality—A Cautionary Note," *Journal of the American Statistical Association* 77(378), pps. 316-319.

Demonstrates some consequences of systematic sampling and temporal aggregation for Geweks's method of decomposing linear relationships between two variables, Y and X, into linear feedback from Y to X, linear feedback from X to Y, and instantaneous feedback between Y and X. Shows that under some conditions systematic sampling will not change the decomposition but it will weaken any relationship between Y and X. In general, however, systematic sampling can manufacture feedback between the two relationships. Temporal aggregation is shown to do this in this article; it also is shown to increase the dominance of instantaneous feedback or contemporaneous correlation between Y and X. Very clear and useful introduction to the mathematical formulation of the two measurement practices. Another important demonstration of the effects that systematic sampling and temporal aggregation have on causal inferences, and on attempts to chart the contours of the 'empirical battlefield' for political theories (viz. the call by proponents of weakly specified models for theories that account for the repeated findings of random walks and high contemporaneous correlations).

(1981) "Effect of Systematic Sampling on ARIMA Models," *Communications in Statistical-Theoretical Mathematics*, A10, pps. 2389-2398.

Corrects some of Brewer's (1973) results about the limiting form of ARIMA models under systematic sampling. In particular shows that ARIMA(p, d, q) becomes approximately an IMA$(d, 1)$, $1 \leq (d-1)$ model as the sampling interval increases. Hence an ARIMA$(p, 1, q)$ becomes approximately a simple random walk. Also derives some results about the speed of convergence to such models. Includes useful mathematical examples for true models of the (2,0,0), (1,1,0), and (0,2,0) types. An important contribution since any general ARIMA model can be approximated by a finite IMA model and therefore the result about the creation of random walks always applies. Should be of interest to users of weakly specified models who take the repeated finding of random walks as objective fact that requires explanation as through rational expectations theories of politics.

(1978) "The Effect of Temporal Aggregation on Parameter Estimation in Distributed Lag Model," *Journal of Econometrics*, 8, pps. 237-246.

Refines and extends Tiao and Wei (1976) and related works on the problem of estimating bivariate, finite distributed lag models. Derives the (true) Direct Aggregate for the underlying model and shows that the aggregated models most researchers actually use again are misspecified and hence likely to yield biased and inconsistent estimates of the true parameters. Shows how Conditional Aggregate can be used to obtain unbiased and consistent estimates of the true parameters. Derives theorem comparing the relative efficiency of the true, Direct Aggregate, Conditional Aggregate models—a theorem which shows that the asymptotic covariance matrices for the estimated coefficient vector is largest for the conditional aggregate model and smallest for the true model. Points out that multicolinearity also is likely to become more severe by virtue of temporal aggregation. Thus it will be very difficult to determine true parameters of model as one aggregates the time series. Includes a useful mathematical example. Somewhat easier to read than Tiao and Wei (1976). Extremely important for users of strongly specified, distributed lag models in that it demonstrates the effects temporal aggregation has on efficiency.

Weiss, A. (1984) "Systematic Sampling and Temporal Aggregation in Time Series Models," *Journal of Econometrics* 26, pps. 271-281.

Summarizes and extends the literature including Brewer's (1973) results. Includes consideration of integrated or ARIMA models and of seasonal models. Shows that seasonality in time series will remain in the systematically sampled and (or) temporally aggregated series but that the frequencies of the resulting cycles will be different as a consequence of the two measurement practices. Concludes with some interesting observations about the likelihood that economists actually have overaggregated their data. Hard to evaluate since, once more, economists have more systematic knowledge about pure character of their time series than political scientists. Brewer (1973) might be read first to gain some understanding of the mathematical reasoning underlying the derivations in this article.

Zellner, A. and C. Montemarquette (1971), "A Study of Some Aspects of Temporal Aggregation Problems in Econometric analysis," *Review of Economics and Statistics* 53, pp. 335-342.

Studies the effects that temporal aggregation has on a simple regression model in the differences into variables. Shows that while temporal aggregation leaves the coefficients of the simple model unbiased and consistent, it produces a loss of efficiency (a biased estimate of the residual variance) and creates a moving average error term. In turn the conventional t and F tests are inappropriate; and, substantive interpretations of the error structure are misleading. Temporal aggregation also inflates the R for this model as a pure "mathematical effect," and create prediction error. Includes illustrations for simple money multiplier models. The analysis in this paper is easy to follow. The practical implications of this analysis may be limited, however. Nonetheless, this probably is the best place to start reading about the consequences of temporal aggregation. TPM

Graphic Displays in Political Science, *Charles H. Franklin, Washington University*

It is well known from the equation

$$\text{Words} = \beta \text{ Pictures}$$

that $\beta \approx 1000$ (with some variance, depending upon the author). With such an efficient conversion process, it is little wonder that scientists find graphical representations attractive. A scatter plot can convey far more information about a relationship than a simple listing of the data, for example. It is little wonder, then, that graphical displays of data are commonly used in scientific writing. It is a bit more surprising that the social sciences use far less graphical material than do the natural sciences.

In a study of 57 scientific publications, Cleveland (1984) sampled 50 articles from each journal during the period 1980-81. He measured the amount of space devoted to graphs as a proportion of total space in the article. The results show that of the journals in the natural sciences, the median was about 11% graphs. The median for 17 social science journals was 2.5%. While 28 of 29 natural science journals used more than 5% graphics, only 2 of 17 in the social sciences were over 5%.[1]

The relative lack of graphical presentations in the social sciences may be due to the brevity of most articles in the natural sciences compared to the relative lengthiness of those in the social sciences. But whatever the reason, the power of graphs to communicate information should not be overlooked by political scientists. In increasing our use of graphical methods, we should also be aware that something as simple as a picture may not be so simple after all. Recent work has shown that some type of graphs convey information better than others. The implications of this work for our use of graphics and our choice of software are the focus of this article.

Lessons for Graph Design

The goal of any graph is to convey information clearly and accurately. Cleveland (1985) sets out a number of flexible rules which help achieve this goal. These suggestions deal with the clarity of the graph, how understandable it is, and the use (or misuse) of scales.

For clear vision, the primary rule is that the data should stand out. The data area of the graph should contain nothing to distract from the point being made. When the data area (that is, the region bounded by the axes of the graph) becomes cluttered with labels, arrows, unnecessary reference lines, or too many lines of data, this objective is lost. Strunk and White apply here: less is more—avoid clutter.

Overlapping data-points tend to confound. If overlap is unavoidable, use symbols which remain visually distinguishable. Open circles can overlap quite a bit yet remain distinctive, for example. If several symbols are used in a graph, then they should be distinct enough to survive some overlap. Various filled and open circles seem to work well, while squares and triangles work less well.

Another common problem concerns the fit of data within the frame of the graph. The data should not touch the frame because this hides the points. If necessary, adjust the range of the axes in order to move the data slightly away from the frame.

An often overlooked consideration is how the graph will look when it is reduced to fit in the journal. A beautiful 8-1/2 by 11 inch graph may be unreadable when reduced to fit a 5 by 6.4 cm column of *APSR*. Cleveland suggests testing on a reducing copy machine—reduce to 66%, then

[1] A check of four randomly selected issues of the *American Political Science Review* from the period 1985-1988, one issue from each year, finds that the average is 4.03%, with a range of 1.78% to 8.81%. That average puts *APSR* slightly ahead of the *American Economic Review* (3.8%) and slightly behind the *British Journal of Psychology* (4.5%). For the *APSR*, articles on political philosophy were excluded from the page count.

reduce the reduction to 66%. If you can still read it, it must be okay.

In order to understand a graph, the reader must be told what is being graphed. One would think this obvious, but after looking at graphs in the last four years of *APSR* I think it is a point worth making. Titles and legends of graphs should be comprehensive and informative. Graphs in the natural sciences frequently have long descriptions as legends. In the social sciences this practice seems far less common, but we could profit from a measured imitation. An example of this concerns the use of error bars. While error bars should be used in many situations, it is crucial that the reader be told if they represent one standard error, two standard errors, a 50% confidence interval or a 95% confidence interval. Omitting this tiny detail prevents clear understanding of the graph.

Some of Cleveland's most useful suggestions have to do with the choice of scales. While it is nice to keep scales in their natural units, this does not always produce a clear picture. Where variables range over several orders of magnitude and may be skewed, plots are likely to be less revealing. For example, a plot of gross domestic product for all nations of the world will suffer from the disparity between the most developed and the less developed nations' economic production. In these cases, the use of a log scale may help. Cleveland points out that if the data range over several magnitudes of 10, then a $\log_{10}$ transformation may be best. If the range is less than this, then a $\log_2$ transformation may be better. The charm of the $\log_2$ scaling is that each unit represents a doubling of the measure, which is relatively easy to understand.

Scales should be chosen so that the data fill as much of the data region as possible. When the data region is largely empty, the resolution is generally low and it is difficult to distinguish points.

Comparisons between graphs are relatively easy if the scales are the same, and relatively impossible if they differ. When juxtaposing graphs, therefore, a premium should be placed on setting the scales equal.

It is very helpful to fully enclose the data area. Many graphs consist of x and y axes only. This makes judgments of values in the upper right quadrant of the graph difficult. Cleveland advises the use of right and top axes to enclose the data area. These right and top axes should have the same tic marks as the left and bottom axes, respectively.

A side benefit of this scheme concerns log scales. If the x-axis is scaled by $\log(x)$, then it may be desirable to label the corresponding tic marks on the top axis with the corresponding values of x. In this way, the log scale helps the resolution of the display, while the numerical values of x along the top help the reader understand the graph in terms of the original scale.

These "principles of graph construction", as Cleveland calls them, provide a useful set of rules to guide graph construction and to critique first drafts of graphs. Like prose, graphs need second (and third) drafts. However, there is more here than meets the eye. Cleveland's work is not simply a set of rules of aesthetic judgment. Experimental evidence on graphical perception shows that there are some sorts of graphs which are "better" than others in the sense that they are more readily understood and less subject to perceptual distortion.

It is possible to order the accuracy of perceptual tasks. To the extent our graphs require easier operations rather than more difficult ones, they will be easier to understand. Cleveland orders these tasks as

1. Position along a common scale
2. Position along identical, nonaligned scales
3. Length
4. Angle and Slope
5. Area
6. Volume

The practical implication is that we would be better off keeping our data points within a single frame so long as they remain distinct. When this is not possible, then juxtaposing graphs with identical scales is the next best thing. Since position is more easily judged than length, we should avoid graphs which require length judgments. An example of such a graph is a stacked bar graph. Here the length of segments of bars represent the values of our variables. But these segment lengths are inherently difficult to judge. Cleveland presents a variety of alternatives to length-based graphs which are easier to read without compromising the portrayal of the data.

The most damaging item in the ordering is the low ranking of area judgments. Experimental evidence shows that areas are commonly misperceived. In particular, the relative sizes of areas are biased so that small areas are perceived as too large while large areas are perceived as too small. Graphs which rely on areas to represent quantities are therefore undesirable.

Implications for Software Selection

Graphical software must be extremely flexible if we are to design graphs with these considerations in mind. Unfortunately, much of the commonly used graphical software lacks this flexibility. Probably the most commonly used graphics software (at least among political scientists) is built into a spreadsheet, such as *Lotus 1-2-3*, *Excel* or *Quattro*. While these packages make it easy to design simple graphs, they do not allow much flexibility in choice of scale, labeling, symbols and legends. While useful for quick and dirty jobs, they are inadequate as general purpose graphics programs.

Several specialized scientific graphics packages provide excellent flexibility in designing graphs. Not only do they provide a good selection of types of graphs (line, bar, box, scatter and so on) but they also allow the user to modify elements of the graph at will. Examples of this genre are *Grapher*, *Sigma-Plot* and *Tech-Graph-Pad*.[2]

Not only do we need flexibility in the graphics package, we also need a good interface with the statistics package. The best possible arrangement, of course, is an integrated statistics and graphics package. *Systat* and *Stata* are examples of well integrated, reasonably flexible combined packages. Of the descendants of mainframe packages, *SPSS/PC+* and *SAS-PC* both offer graphics capabilities as well as an integrated data analysis package. These last two are, however, quite large and will eat anything less than a 40M or even 60M hard disk.

Based on an unsystematic review of 8 graphics systems, it seems to me that the graphics only packages suffer from the inevitable complications involved in moving data from the statistical to the graphics package. While they can produce very pretty graphs, the inability to play with the data while iteratively producing a graph is a serious limitation. Of the integrated statistics and graphics packages, I've found *Systat* to be the most comprehensive, both graphically and statistically, though at a rather steep price. *Stata* is quite inexpensive at the academic price, while offering a nice, reasonably flexible set of graphics. *Stata* is not, however, nearly so complete a statistics package as *Systat*. The other integrated package I've used, *SAS-PC*, is too large, cumbersome and slow for my tastes, though it has many of the advantages of mainframe *SAS*. Finally, *Gauss* has a set of presentation graphics routines which are very flexible. If you have already climbed the steep *Gauss* learning curve, these graphics routines may suit your needs well. And of course, in *Gauss* you can always write any extensions you need.

References

Cleveland, William S. 1984. "Graphs in Scientific Publications", *The American Statistician*, **38**: 261-269.

Cleveland, William S. 1985. *The Elements of Graphing Data*, Monterey, California: Wadsworth.

PC Magazine 8: 5 (March 14, 1989) pp. 121–285. TPM

[2]For a review of these and other graphics packages, see *PC Magazine*, vol. 8, no. 5 (March 14, 1989). The same issue also reviews a number of statistical packages, including those mentioned below.

The 1989 Political Methodology Meetings, *Larry Bartels, University of Rochester*

The University of Minnesota hosted the Sixth Annual Political Methodology Meetings in Minneapolis July 13-15, 1989. The meeting was organized by John Freeman for the host institution and by Stanley Feldman for the Political Methodology Organized Section of APSA. The meeting was somewhat larger than in previous years, with 31 scholars from across the country among the invitees. Fourteen of these 31 had not attended any of the five previous annual meetings; seven were graduate students.

Presentations at the meeting's six sessions varied widely in content and style. As at previous meetings, questions and comments from the floor were frequent, spirited, and wide-ranging. The presenters and topics included:

Christopher Achen, University of Chicago (Prospective Voting and The Theory of Party Identification)

James Stimson, University of Iowa (Measuring Public Policy Mood)

Charles Franklin, Washington University (Estimation Across Datasets: Two-stage Auxiliary Instrumental Variables)

Gary King, Harvard University (Some Thoughts on Ecological Inference)

Elisabeth Gerber and John Jackson, University of Michigan (What If Institutions and Preferences are Endogenous?)

Henry Brady, University of Chicago (Statistical Consistency for Metric Multidimensional Scaling for One Person Who Compares Many Candidates), Discussant: Douglas Rivers.

Moderator on Time Series Models: Nathaniel Beck, University of California, San Diego

Mack Shelly, Iowa State University (Federal Budgetary Statistics)

Walter Mebane, University of Michigan and Cornell University

Michael Mackuen, University of missouri–St. Louis (Roundtable on Modeling Time Series Data)

Philip Schrodt, University of Kansas (Inductive Modeling of Event Sequences Using Computational Methods)

Douglas Rivers, UCLA and Stanford University (Selection Bias in Linear Regression, Logit, and Probit Models)

Larry Bartels, University of Rochester (Squared Error Analysis of The Instrumental Variables Estimator Under Misspecification)

William Berry, University of Kentucky (Using Event History Analysis to Study State Policy Innovation) Discussant: Michael Goldfield, Cornell University

Additional participants included R. Michael Alvarez, Duke University; Stephen Ansolabehere, Harvard University and UCLA; John Brehm, University of Michigan; Patricia

Conley, University of Chicago; John Freeman, University of Minnesota; Michael Hollis, UCLA; Herbert Kritzer, University of Wisconsin; Keith Kybee, University of California, San Diego; Christopher Mooney, University of Wisconsin; John Williams, University of Illinois at Chicago; and Frank Zinni, Richmond, Virginia. Several faculty members and students from the University of Minnesota also attended some or all of the sessions.

The meeting was financed by the University of Minnesota and by a grant from the National Science Foundation. Previous political methodology meetings have been held at UCLA (1988), Duke University (1987), Harvard University (1986), The University of California, Berkeley (1985), and the University of Michigan (1984). Washington University in St. Louis is scheduled to host the 1990 political methodology meeting in St. Louis. TPM

"Creating" Good Problem Sets,

John E. Jackson, The University of Michigan

The Problem of the Problem Set

Arranging good problem sets is the most vexing part of teaching any methodology course. Texts and articles demonstrating both good and less than adequate applications of different techniques are selected without too much difficulty. Problem sets are another matter.

The first step is to scower the existing well documented datasets in the ICPSR catalogue, many of which are probably already available locally. This excursion reveals several discouraging facts. The substantive content of many existing data sets is too complicated to be neatly excerpted for a methodology problem set. Imagine a simple four variable model of political attitudes or growth in the public sector. In other instances, because of the "real world" nature of the data, the particular problem to be identified and treated is intertwined with other statistical problems that greatly complicate the analysis, and may even dominate the problem being illustrated. Getting students to diagnose and treat autocorrelation in circumstances that also present problems with limited dependent variables, with errors in variables, or with simultaneity is a daunting task, much like trying to replace one fixture in an old house. (A circumstance that led to a $10,000 disposal in one former house.) At this point, one usually makes do by simplifying what is in the data and telling the students to ignore other estimation problems. None of this seems very satisfactory.

There is another way to "fake it." I decided several years ago in an advanced methods course to create my own problems and datasets using Monte Carlo simulation. This approach has a number of nice features, some of which were not evident until after students began their analysis. This article describes some of these advantages, the procedures used, and offers one example taken from the advanced methods course.

Creating a Problem Set

The concept of the Monte Carlo simulation is very simple. (See E. A. Hanushek and J. E. Jackson, *Statistical Methods for Social Scientists*, New York: Academic Press, 1977, for repeated pedogogical uses of these simulations.) The problem set is defined by the substantive scenario used to frame the exercise and by the particular parameters chosen by the instructor, such as the number of variables and observations. Random numbers are then drawn to represent the stochastic elements in the model and included in this structure. The specification of the random numbers and the way they are included in the structure are determined by the problem being assigned. For example, if one is simulating autocorrelation, the stochastic term at time t is a function of the stochastic element at time $t-1$ and a random term, $u_t = \rho u_{t-1} + e_t$. This term is included in the structure in an additive manner. If an errors in explanatory variables problem is assigned, the random number drawn is added to the relevent exogenous variables, $X_{tk} = X_{tk} + e_t$. Separate datasets can be created using the same values for the exogenous variables and different draws from the distribution of random numbers. In this way, each student gets a separate "problem" set to analyze.

The key to the problem set is the scenario that describes the substantive arguments, the hypotheized model, the data, and the conditions that lead to an estimation problem. Often, these scenarios build on articles read as part of the assignments. For example, the early assignments in the methods course include several papers modeling growth in government expenditures in democratic countries. The first problem set describes two data sets measuring public sector expenditures and a set of exogenous variables reflecting the core hypotheses. The variables are the same in each set. One dataset represents cross sectional data collected from a set of forty countries while the second set represents data taken from a single country over a forty year period. Students are asked to estimate and compare the relationship between various factors and the size of government with these two data sets. It is suggested that the cross section data may present heteroscedastic problems, with the variance of the deviation of actual expenditures from expected expenditures being positively related to national income. They are also told to expect serial correlation problems with the time series data. Students are asked to select and justify an estimation procedure appropriate for each data set and to compare methods if several methods are used.

There are several advantages to these simulated datasets. The most obvious one is that the instructor can make sure that the condition to be diagnosed and overcome exists within the data. For example, in the above examples, the simulations are designed so that heteroscedascticity and autocorrelation are very evident in the data. Conversely, the data are developed so that other possible problems are minimal or non-existent. These conditions, though idealized and not characteristic of real data, are very useful in helping students to recognize the problems at issue and to see the effects of different estimation methods. My experience is that students learn better when their first encounter with situations are clear and dramatic. It is important for them to go, "Oh! So that's what autocorrelation looks like," and, "Wow! GLS makes a difference in the standard errors but not much in the estimated coefficients." I have found that it is harder to get these revelations of truth with real datasets.

Some of the less obvious advantages turn out to be even more important. We usually work with, and teach, the assumption that any dataset is just one possible draw from a distribution of many possible outcomes.@foot:Even if we have all years in some time period or all cross section units, such as states or legislators, we are still sampling behavior. The observed outcomes are not the only possible outcomes for that year, that state, or that legislator. (Brady offered a nice discussion of this assumption in the previous newsletter.) Consequently, the results obtained from analyzing any particular dataset are unique to those data and will not equal the true structure presumed to generate the data. Our statistical analysis is designed to allow us to say something about the distribution of possible results, given the distribution of possible datasets, and how that distribution relates to the presumed structure that generated the data.

I have found this concept of a dataset and the associated results being simply one draw from a distribution to be very difficult concept to grasp, and not just for students. People often see one dataset and the results obtained from analyzing those data and presume they have the answer to the question of how certain variables relate to each other. The Monte Carlo simulations offer as many different datasets as their are students, more if the instructor wants to pursue the matter. Consequently, when students compare results, as they surely do, and when it comes time to discuss the problem in class, people discover a wide range of "answers." (This lesson is aided by tabulating estimates for each coefficient on a blackboard.) Even more revealing is the explanation that each of these answers is correct. Correct in the sense that the proper procedures were followed and correct given the particular dataset.

The distribution of student results produces a second important lesson. Revealing the "true" values used in the simulation are is necessary at this point. (If this is the first problem set, Monte Carlo simulation needs to be explained.) Some—many—of the estimated values will differ substantially from the true values. The pedagogically useful ones are those with a different sign from the true value, with extra points given for the ones of these that are statistically "significant." Students emerge from this demonstration, particularly after several exercises, with a healthy scepticism for what truth about social processes is contained in one set of results. At this point, it is again useful to underscore that fact that, we hope, everyone did their statistical analysis correctly. The deviations of the results from the true structure is not attributable to incorrect method or technique, but to the vagaries of data.

Invariably, at this point someone asks, "Then, what can be learned from a dataset!?" My answer is that empirical research is a game played against nature. The objective of empirical work, the game, is to make the best possible guess about the natural structure that generated the data measuring the behavior of interest. Nature, of course, keeps the information about this structure secret. The hypothesized model, or models, are the initial guesses about this structure and the statistical analysis indicates how "good" these guesses are, as assessed by their consistency with the data.

This game against nature must be played according to very specific rules by people who are patient and clever. (Remember the advertisement that, "It is not nice to fool Mother Nature.") I have found that most students develop a better appreciation for the purpose of empirical work and for the complexity of the task after seeing the distribution of results and how they differ from the systematic part of the data generation process. It is also evident at this point that statistical technique and data alone are not sufficient to play the game successfully. One must always be making guesses and checking them against the data. After all, students have direct evidence that some "correct" analysis just led to results that are completely contradictory to the structure nature used.

The answer just outlined also establishes the instructor as a very important figure for the rest of the class. The person controlling the simulations is, in effect, nature and the students are trying to "guess" what structure nature has used. This is a social scientist's one (only?) chance to play God in some aspect of the world. (We may be doctors, but we are not physicians.)

The example exercise based on simulated datasets that I offer illustrates this notion of the students playing a game against nature. The section on limited dependent variables presents alternative ways of conceptualizing the responses to simple trichotomous choice questions, such as the agree, disagree, no opinion/don't know found in survey research. One option is that the responses constitute an ordinal assessment of preferences with no opinon/don't know constituting indifference. In this case, the responses are best modeled by an ordered probit model. Secondly, the no opinion/DKs may be a completely separate category of response that is quite distinct from preferences. In this case, the responses constitute a categorical variable which may be modeled by a multinomial logit expression. Lastly, we discuss a decision tree type model. In this model, people first decide if they have an opinion when asked the question. If they decide they have an opinion, then they indicate whether it is an agree or disagree. In this model, we first model the probability a person has an opinion. Next, we model the probabilities people agree, conditional on their having an opinion. Each of these models implies a different statistical estimator and test for consistency with the data.

The problem set presented below presents students with data about preferences commonly found in survey research. They are asked to guess the process underlying this choice behavior and to estimate the effect of individual differences on the likelihood of opposing, favoring, or having no opinion. The estimate of these individual differences varies considerably with the model hypothesized and selected by the student. Each student is given a dataset with 300 simulated responses and corresponding values for the explanatory variables. The simulated responses were generated with the conditional choice model described above, with an expected no opinion/DK response from 20% of the sample.

An Example

In estimating a simple model of attitudes towards increased local school expenditures, the following variables were constructed from a survey of residents:

Y = Attitude (1 = Oppose, 2 = Favor, 3 = No Opinion),

X_1 = Education in years of schooling,

X_2 = Black,

X_3 = Age in years,

X_4 = School age children in family,

X_5 = Income in thousands of dollars,

X_6 = School age children in non-public school, and

X_7 = Employed as teacher.

Compare the probability of having no opinion or of favoring increased expenditures of a 28 year old black, with 18 years of education, earning $50,000, who has a child in public schools and who is not a teacher with that of a childless, 65 year old white, with 12 years of education, earning $20,000 and who also is not a teacher. What is the marginal effect on these probabilities if schooling were to decrease by a year?

There is considerable discussion in the methodological literature about how to treat the no opinion response in survey research. Does this response indicate indifference, i.e. an intermediate point on the oppose-favor continuum? Or is it indicating the absence of an attitude, which constitutes a separate category, that is not on the continuum? How would you treat each of these alternatives in answering the above questions?

Student Reaction

Student's adapted very quickly to the idea that they were not dealing with "real" data. I had been concerned that they might be unable to suspend belief in a way that would permit them to take the problem sets seriously. This turned out to be no problem. Maybe graduate students are better than faculty at suspending belief, at least about certain things. They take the model and data in the problem so seriously that I regularly get discussions of omitted variables, possible simultaneity among variables, and other "real world" comments.

Students seem to get very caught up in the game against nature aspect of the problem sets. In a friendly and constructive way, they compete with each other to see you comes the closest to getting the correct specification. They also see it as a game against the instructor to guess the right model. These games within games add interest and get their attention. I also observe a lot of exploration of different hypotheses, models, and even other techniques as students try to figure out what nature gave them.

My feedback has been that they find the problem sets one of the most valuable parts of the course. I am concerned about what that says for the readings and lectures, but it does indicate success in the applied parts of the methods course. TPM

Choosing a Computer Package for Political Methodology Courses,

Nathaniel Beck, University of California, San Diego

I teach three types of methodology courses: a lower division statistics course required of all sophomores majoring in the social sciences (including economics) and the second and third course in our graduate political methodology sequence (the third course is also taught at a stand alone at the ECPR summer school at Essex). All three courses require student use of the computer to conduct data analysis. I thought it might be helpful to others if I went through some of the issues that arose in my choice of computer package for the three courses.

The choice for the lower division course was heavily constrained. We do not have micro labs suitable for the 300 students per quarter that we teach and the university would like all undergraduates to be familiar with UNIX. We therefore had to find a good general purpose package that could service this many students in a UNIX environment without dramatic degradation of performance on a minicomputer. Many of our students have never seen a computer before, so we needed a package that was easy to use and well-documented. Some students taking the course would go on to more advanced courses in sociology, political science and economics, so we needed a package that contained a variety of powerful procedures useful in the various social sciences. (My choice would have been different if I were teaching the course in the political science department.)

Given these desiderata, we chose to use BLSS, the Berkeley Interactive Statistical System (formerly known as Berkeley ISP). BLSS runs well in a UNIX environment, has excellent documentation (the Abrahams and Rizzardi Blss Book, published by Norton), works well with the course text (Freedman, et al.), and is the package used in the introductory econometrics course (to which my course is prerequisite). BLSS is not quite as good for standard political science problems as some other packages (e.g. SST) but it is quite decent for those problems, and also contains a variety of powerful procedures useful in a wide variety of contexts.

Most of the problems I had with BLSS were really problems with UNIX. Students would prefer the least user friendly package running under DOS to the most user friendly package running under UNIX. But here I had no choice. Once students got past UNIX problems (e.g. redirection, local printing, unreadable documentation), they found BLSS reasonably easy to use. The language of BLSS is consistent with the language of Freedman, et al, and students could use social science datasets to illustrate what they were learning in Freedman. The BLSS manual is quite readable, and the on-line help is quite good. (My one quibble with using BLSS at this level is that some of the recode type commands seem quite unintuitive and the student has to deal with missing data before running a regression. There were a few days that I longed for SPSS, but only a few.)

The two graduate courses use our departments microcomputer lab and a minicomputer (again UNIX) devoted to research in the social sciences. Because of the lack of resources in our lab (small number of non-networked machines), I wanted a package that would run both in the micro and mini environment. I also wanted a package that was relatively easy to use, so students would not have to suffer learning an unfriendly computer language at the same time they were learning about matrices, etc. Most empirical research at UCSD is econometrically oriented, so I needed a package that was similarly oriented. Finally, my course serves as a training ground for our RA's, and most of our faculty use SST in their research.

The choice of SST (Dubin/Rivers Research) for the first graduate course (at the level of Kmenta) was thus fairly easy. SST has worked out quite well. Several of my students got RAships because they knew SST. The class has no difficulty doing interesting regressions, and little difficulty in dealing with large data sets (subsets of major election studies). (The only problem with analyzing the large data sets is that students must use the UNIX machine for those analyses, and they find UNIX no less inpenetrable than do my sophomores.) SST does a fabulous job with its regression command, and so most of the things I teach can be illustrated with SST. SST also has specialized procedures for most of the special topics I teach (primarily limited dependent variables). (SST lacks reasonable time series capabilities, but that is only a problem for a bit of this course.)

The second graduate course (and the ECPR course) is a special topics course using King's *Unifying Political Methodology* (Cambridge). The theme of the course is that the student can estimate any model of interest, mainly via maximum likelihood. The topics are primarily cross-sectional. What SST does, it does very well, but its maximum likelihood routine is not very flexible, and it implements some procedures (e.g. count and duration models) differently than I teach them. Because SST is a package, there really is no facility for me to rewrite their count and duration modules in a way I would prefer.

I thus chose to use GAUSS (Aptech Systems) for this course. Last year this decision lead to a disaster. But after rethinking things, I decided to go with GAUSS again this year. Last year I just turned students loose on GAUSS. It was too much for them. I thought they would be able to use the flexibility of GAUSS to do all sorts of things that interested them; instead they found GAUSS so hard that they could barely do the things that I showed them exactly how to do.

Last year I simply had students use a wide variety of the GAUSS written procedures to do such things as selection bias, switching regression, etc. (Gary King also provided me his count data programs.) The problem was that each procedure was slightly different, and the students became frustrated, spending all their time dealing with minor problems, and differences (e.g some routines automatically included constant terms, others didn't; some automatically printed results, others required print commands to be added).

This year there is a new version of GAUSS. Most of the early procedures have been scrapped. There is now a consistent way of writing GAUSS maximum likelihood procedures. Aptech now provides a set of consistently written (and excellent) limited dependent variable programs, and Gary King has rewritten his count data and duration programs to be consistent with this scheme. I have also rewritten a few of my programs (selection bias and switching regressions) similarly. I hope that by providing the student with a set of modules with identical calling procedures I can avoid teaching any more than five minutes worth of GAUSS.

The mistake I made last year was to try to go without a package. This year I am thinking of GAUSS as giving me the ability to write my own customized package. This takes more work on my part than would the use of SST, but, given the hard work done by Aptech and Gary King, designing a personalized package with GAUSS is really not very difficult (remember, I already know GAUSS!). Unfortunately, I haven't tried this scheme out in practice. I shall know more after this summer. But for now my feeling is, if you want to use a prewritten package, SST seems best, but if you want flexibility GAUSS offers that (at a price).

What about time series? SST doesn't really do them, and GAUSS doesn't have a good set of procedures yet (although I do have my Kalman filter routines). The last time I taught time series in depth I used RATS (VAR Econometrics). I actually use and like RATS, but my students find it hard. It is, however, one of the only micro based programs that allows the user to undertake a variety of different types of analyses, and the new version (3.0) is friendlier than previous versions (is this damning with faint praise?).

Finally, I would like to thank Dubin/Rivers, Aptech and VAR for allowing me to use their programs at Essex without paying a license fee (which the ECPR could not afford). I have found the authors of all the four pieces of software mentioned in this article to have been most helpful, and quite responsive to my needs, as well as willing to listen to my myriad criticisms. TPM

A Statistics Syllabus for Undergraduates: Political Science and Sociology, *George E. Marcus, Williams College*

COURSE CREDO
I hear, and I forget;
I see, and I remember;
I *do*, and I *understand.*
-Chinese proverb

Course Readings, Required for purchase

1. George W. Bohrnstedt and David Knoke, *Statistics for Social Data Analysis*, 2nd Edition (Peacock,1988)
2. James A. Davis, *The Logic of Causal Order* (Sage, 1985)
3. *Political Science 206 Documents* Packet of supplementary materials and readings sold in Political Science Department, Stetson. A list of course documents is at the end of the syllabus.

The reading load is light. This is deliberate; it represents an attempt to combat the well known tendency for a kind of Gresham's law to operate in liberal arts colleges: the more important activity, thinking, is driven out by the less important activity, reading.

Students experiencing problems, i.e., either having difficulty in understanding the material or finding sufficient intellectual challenge in the material, are urged to consult the instructor.

Bring the textbook and relevant documents/readings to class, as the instructor will refer to specific passages from time to time.

Course Requirements

1. The course consists of the following activities (Failure to complete the first three will adversely affect course grade):
   (a) Chapter Problems or Exercises, P/F
   (b) Individual Assignments, P/F
   (c) Group Exercise, P/F
   (d) Analysis Paper, 60%
   (e) Mid-term Examination, 25%; The mid-term grade will be included if and only if it increases the final grade
   (f) Final Examination, 40%
   (g) Class Participation, (rounding factor)
2. Many problems or exercises at the end of each chapter of Bohrnstedt & Knoke are well worth doing. That is, doing them will help you learn the material. Throughout the syllabus you will find listed some of the exercises we find particularly worthy of your attention. Unless otherwise specified, these are not to be considered written assignments. Some, however, have been specifically assigned to be turned in and graded on a pass/fail basis.
3. The group exercise and the individual exercises involve the use of the Macintosh microcomputer. No knowledge of programming is necessary; students are taught how to use political science software programs.
4. The examinations require interpretation of data, mastery of concepts, and evaluation of specific theories, rather than extended, synthetic essays.
5. The Analysis Paper of approximately 15-20 pages involves the use of data sets. A detailed write-up is provided in CD-11, Analysis Paper. The Analysis Paper can and *should* be started as soon as possible after the mid-term, in order to avoid a serious end-of-the-semester crunch.
6. Much of the reading material in the course is somewhat technical; but the number of pages is small. Most students will find they will learn more efficiently if they carefully complete and review the readings before class, make use of the suggested problems, listen attentively in class, to make certain that thorough understanding has been attained. In the examinations, the *use* of the statistical concepts is stressed, rather than memorization of formulae, the calculation of coefficients, or the mastery of mathematical derivations. Nevertheless, the instructor holds the view that "understanding" requires a substantial mastery of statistical concepts.

7. For your convenience the mimeographed handouts have been numbered with either a "CD" (Course Document) or "CR" (Course Reading) prefix.
8. The weekly schedule provides for two lecture meetings per week. While questions are welcome during and after each lecture, additional opportunity to deal more intimately with the materials and subject of the course will be provided by the laboratory sessions. In addition, lab sessions will be used to present some additional subjects and grasp the group exercise.

Course Evaluation

Student opinions concerning how the course is going, e.g., what has been left unclear, whether readings and exercises are meeting their objectives, etc., are valued, and students are encouraged to express their views *throughout* the semester to the instructor and to the student assistant(s). The Student Course Survey is given at the end of the course. Student opinions are considered seriously each time the course is reviewed for possible changes.

Student Assistants

The student assistants have been appointed to provide you with assistance in a wide variety of course activities: using the Computer Center, formulating your hypotheses and interpreting data for exercises and the analysis papers, understanding the statistical concepts, etc. Scheduled hours are kept to a minimum, in order to maximize their availability when exercises are due and for individual appointments. Feel free to call the student assistants for individual appointments throughout the semester.

The students assistants are *not* intended to limit access by students to the instructor outside class, but rather to supplement such access. Feel free to contact your instructor concerning any aspect of the course throughout the semester.

Academic Honesty

All provisions covering academic honesty and the Williams College Honor Code, described in detail in the Williams College STUDENT HANDBOOK, are of course in force for all aspects of student work in Political Science 206. Except the Group Exercise, students are expected to do their own work on all exercises and problems. On the group exercise students work cooperatively and each student is expected to do his/her fair share of the group's work.

The two examinations are scheduled and are held in an classroom, where each student is expected to present and formulate his/her own answers without consulting others and without reference to any materials other than the examination itself.

For the analysis paper, which is to be done individually, the student *ought* to consult with the instructor, the student assistants, classmates and other helpful sources, both written and oral. Academic honesty requires that ideas and information the student receives from any of these sources be acknowledged appropriately. *Should the student make use of an Analysis Paper written by another student during a previous semester of Political Science 206F or 206, a copy of that Analysis Paper must be attached to his/her own paper when it is submitted.*

Weekly Reading Assignments

Meeting 1 The course begins with an examination of what constitutes theory and what constitutes a persuasive explanation in social science. Key concept: Property Space.

Meeting 2 I. Introduction: Course Rationale and Organization CD-1 Syllabus: CD-2 *Creating Explanations*, Chapter 1 - Introduction; CD-3 *Creating Explanations*, Chapter 2 - Explanation CR-1 Muller and Seligson, "Inequality and Insurgency." Supplimental Reading (Optional): B&K, ch. 1. LAB #1: Using the Macintosh Computer

[The next topics we consider are categorization and measurement. Key concepts: values, attributes.]

Meeting 3 II. Categorization and Measurement: Continuous and Categorical Classification. CD-4 *Creating Explanations*, Chapter 6 - Measurement Basics. *Required:* Exercises - all exercises in CD-4

Meeting 4 III. Measurement: Measuring Single Variables; Bohrnstedt & Knoke Chapter 2; *Required*: B&K Chapter 2, problems 4, 5, 9, 19; Using Data Desk Professional; LAB #2: Using the Macintosh Computer and Data Desk Professional CD-5 Using Data Desk Professional

Meeting 5 Bohrnstedt & Knoke, Chapter 3, pp. 65-75, 80-91; *Required*: B&K Chapter 3, problems 1, 2, 5, 15, 16, 19, 21. LAB #3: Using Data Desk Professional (Cont'd); CD-6 Individual Assignment 1; CD-7 Individual Assignment 2.

[The next topic we turn to is covariation. Evaluating events and attributing cause and effect are essential steps in achieving understanding. For a cause to explain an effect we need to demonstrate that they are linked - that they covary. Thus, covariation is at the heart of all kinds of explanations.]

Meeting 6 IV. Measurement: Descriptive Statistics - Bivariate Relationships; Bohrnstedt & Knoke, Chapter 8, pp. 256-258

Meeting 7 Bivariate Relationships - regression and correlation Bohrnstedt & Knoke, Chapter 8, pp. 253-275

[The next topic we turn to is sampling. We normally do not study entire populations. Rather we observe a selected subset of the population of interest. When can we generalize what we find in the subset to the population of interest?]

Meeting 8 V. Statistical Inference: Bivariate Relationships; CD-8 *Creating Explanations*, Chapter 7 - Sampling Theory

Meeting 9 Bohrnstedt & Knoke, Chapter 5, pp. 144-178; *Required*: B&K, Chapter 5, problem 12; CD-7 Individual Assignment 2.

Meeting 10 Bohrnstedt & Knoke, Chapter 8, pp. 277-289, 296-297; *Required*: B&K, Chapter 8, problems 4, 30; CD-6 *Individual Assignment 1 due.*

Meeting 11 Experiments: Cause and Effect - The ANOVA Model; Bohrnstedt & Knoke, Chapter 7, pp. 219-236

Meeting 12 MID-TERM EXAMINATION *Covers material through Meeting 11 90 minute examination - will be scheduled for an evening examination. CD-7 Individual Assignment 2 due* (bring to exam).

[Spring Break: March 17th to April 2nd]

[The next topic is an extension of the earlier subject, covariation. Explanations may include more than one cause for a specific effect. How we can explore and examine whether there are multiple causes is the concern of multivariate analysis.]

Meeting 13 IV. Multivariate Relationships - Causal Models; Davis, Chapter 1, pp. 7-34; Bohrnstedt & Knoke, Chapter 11, pp. 381-398; (except pp. 384-385). LAB #4: Mississippi Data and Group Exercises; Introduce Group Exercise: Form Groups; CD-9 Group Exercise CD-10 Mississippi Data Set

Meeting 14 Bohrnstedt & Knoke, Chapter 11, pp. 399-415 CD-11 Analysis Paper

Meeting 15 VII. Multivariate Causal Analysis; Davis, Chapters 2-4, pp. 34-69; Bohrnstedt & Knoke, Chapter 12; CD-12 Project Proposal Description. LAB #5 Mississippi Data Group Exercise Presentations.

Meeting 16 *Path Analysis* Using Data Desk Professional to perform PATH Analysis

[Typically assigning numbers to qualities is not a perfect operation. How adequate (either as a matter of precision or as a matter of categorization) will depend on a variety of factors. More importantly, the degree of success may vary considerably from project to project. Scaling is a method of improving the accuracy of measurement and also a method for assessing the reliability and validity of measurement.]

Meeting 17 VIII. Measurement Theory: Multiple Indicators and Scaling; Bohrnstedt & Knoke, Chapter 11, pp. 382-386; CD-14 *Creating Explanations*, Chapter 12 - Multiple Indicators and Scaling; *Project Proposals due* (see CD-12). LAB #6: Scaling with Data Desk Professional; Introduce Individual Assignment 3; CD-13 Individual Assignment 3.

Meeting 18 CD-4 Creating Explanations, Chapter 6 - Measurement Basics; CR-2 Bollen & Barb, "Pearson's R and Coarsely Categorized Measures"

Meeting 19 caling Applications; CR-3 Bollen, "Issues in the Comparative Measurement of Political Democracy"; CR-4 Bollen, "Political Democracy and the Timing of Development"

Meeting 20 Path Example. CR-5 Smart, "College Effects on Occupational Status Attainment." *Individual Assignment 3 due*

[Some times the data available is not amendable to the requirements of continuous measurement. There are bivariate and multivariate statistical methods for evaluating theories using categorical data. A brief introduction and assessment of some of these methods constitute the final section of the course]

Meeting 21 Bivariate Categorical Analysis; Bohrnstedt & Knoke, Chapter 4; Chapter 9, pp. 305-326; 328-334

Meeting 22 Multivariate Categorical Analysis; Bohrnstedt & Knoke, Chapter 10, pp. 349-371; CR-6 Glock, et al, "Adolescent Prejudice." Individual Assignment 3 returned at professor's office

Meeting 23 no class - work on analysis papers; Individual review of analysis projects with the professor (to be scheduled)

Meeting 24 *Summary and Course Review.* Student Course Survey forms and Supplementary. Questionnaire to be filled out by students

Submission of Analysis Papers

All papers are due at 5:00 p.m. on May 12th, the official ending of classes. Permission to submit Analysis Papers after this date *cannot* be granted by the instructor, but only by one of the Deans. This is a college regulation.

Final Examination

1. Regularly scheduled examination: date, time, and place to be announced.
2. Exam is two and one half hours.
3. Includes all course materials, but with greater emphasis placed on materials since the Mid-term.

Course Documents

CD-1 Syllabus

CD-2 Marcus, Sullivan and booth *Creating Explanations*, Chapter 1-Introduction

CD-3 Marcus, Sullivan and Booth *Creating Explanations*, Chapter 2-Explanation

CD-4 Marcus, Sullivan and Booth *Creating Explanations*, Chapter 6-Measurement Basics

CD-5 Using Data Desk Professional

CD-6 Individual Assignment 1 - Frequencies using the Contemporary Democracies Study

CD-7 Individual Assignment 2 - Scatterplots and Cross-tabulations using the Contemporary Democracies Study

CD-8 Marcus, Sullivan and Booth *Creating Explanations*, Chapter 7-Sampling Theory

CD-9 Group Exercise

CD-10 Mississippi Data Set

CD-11 Analysis Paper (include Steps in doing Analysis Paper)

CD-12 Project Proposal Description

CD-13 Individual Assignment 3

CD-14 Marcus, Sullivan and Booth *Creating Explanations*, Chapter 12-Multiple Indicators and Scaling

Course Readings

CR-1 Muller and Seligson, "Inequality and Insurgency"

CR-2 Bollen & Barb, "Pearson's R and Coarsely Categorized Measures"

CR-3 Bollen, "Issues in the Comparative Measurement of Political Democracy"

CR-4 Bollen, "Political Democracy and the Timing of Development"

CR-5 Smart, "College Effects on Occupational Status Attainment."

CR-6 Glock, et al, "Adolescent Prejudice." TPM

An ICPSR Summer Session Syllabus and Annotated Bibliography on Artificial Intelligence, *Philip A. Schrodt, University of Kansas*

This course will consist of lecture in the morning and laboratory sessions in the afternoon. The lectures will introduce general concepts; in the afternoon we will have the opportunity to experiment with these techniques using a variety of programs. The course assumes familiarity with computer programming, preferably Pascal.

Monday: Introduction to AI in the Social Sciences

Lecture:

1. What is AI and how might it be relevant to the social sciences; how our problems differ from their problems; classes of formal models and formal knowledge representation in social science modeling.
2. The fires beneath the smoke: what AI can and cannot be expected to do at the moment.
3. Historical and contemporary organization of research in mainsteream AI.

Lab:

1. Modeling your problems – a survey of the problems and processes the class is interested in working on using AI techniques. We will be working with these during the remainder of the week. No hiding! This session will involve some introduction to mathematical modeling as well as to AI.
2. Brief review of Pascal and the use of Turbo Pascal on MS-DOS computers.

Tuesday: Rule-Based and Expert Systems

Lecture:

A. Human-coded rule-based systems (expert systems)

1. Rule-based knowledge representation; expert systems; classification problems
2. A very simple rule-based system: the binary tree.
3. "Knowledge engineering" and related issues

B. Machine-coded systems

1. CLS/ID3
2. AQ systems

Lab:

1. Implementing a simple binary tree in Pascal.
2. Two "commercial" expert systems (National Collegiate Software Clearinghouse)
3. The CLS machine-learning system
4. A simple AQ machine-learning system

Wednesday: Self-Organizing Systems

Lecture

1. Genetic algorithms. a. Holland classifiers b. Genetic algorithms for the solution of games and planning problems
2. Neural network and connectionist systems.

Lab:

1. HCLASS: A simple Holland classifier
2. Neural network simulation in Pascal

Thursday: Simulation and Problem-Solving

Lecture:

1. Production systems, blackboard systems and related techniques
2. Declarative systems; an introduction to PROLOG

Lab:

1. Implementing a simple production system simulation in Pascal
2. Implementing simple problem-solving simulations in PROLOG

Friday: Language, Pattern Recognition and Sequence Analysis

Lecture:

1. Natural language. a. Formal grammars and the formal analysis of natural language. b. Cheap and dirty tricks for the processing of natural language. c. Query-by-example techniques
2. Sequences. a. Substitutable sequence comparison: Levenschtein metrics. b. Partially-ordered sequences: Heise structures and production systems. c. Non-linguistic grammars; finite-state machines; syntactical pattern recognition.

Lab:

1. Heise's ETHNO program for analyzing partially ordered sequences
2. Garson's WordMatch automated content analysis

Selected Bibliography of Artificial Intelligence Literature Relevant to the Social Sciences

As noted in lecture, the main problem in applying AI to the social sciences is that problems which are most interesting to us are frequently marginal to mainstream AI and vice versa. Ergo at the moment one has to dig through a lot of irrelevant material to find things of use: it is rather like trying to find the Richardson model and Arrow's Theorem in standard mathematics texts. Neverthless, the books listed below will provide some guide to learning about AI while avoiding books on robot vision systems... For obvious reasons, this is heavily oriented towards applications in political science – I would appreciate suggestions concerning additional literature.

General Surveys of AI and Background

Paul A. Anderson and Stuart Thorson. 1982. "Artificial Intelligence Based Simulations of Foreign Policy Decision-Making." *Behavioral Science.* 176-193.

Stephen J. Andriole and Gerald W. Hopple. 1988. *Defense Applications of Artificial Intelligence.* Lexington MA: Lexington. Focuses on the DARPA AI initiative – killer robots and battle management software.

Aaron Barr, Paul Cohen, and Edward A. Feigenbaum. 1982. *The Handbook of Artificial Intelligence.* Los Altos, CA: William Kaufmann. A bit dated and not worth buying at full price (about $120) but is frequently available at a very reduced price.

Eugene Charnick and Drew McDermott. 1985. *Introduction to Artificial Intelligence.* Reading, MA: Addison-Wesley. Thorough introduction; LISP oriented

Stephen Cimbala. 1987. *Artificial Intelligence and National Security.* Lexington, MA: Lexington Books. Deals mostly with forecasting and expert systems; political science and foreign policy orientation; no killer robots.

Douglas Hofstader. 1980. *Godel, Escher, Bach: Eternal Golden Braid.* New York: Vintage Books. (Recreational epistemology)

Allan Newell and Herbert Simon. 1972. *Human Problem Solving.* Englewood Cliffs: Prentice- Hall. This is classic in terms of formally modeling human decision-making.

Nils J. Nilsson. 1971. *Problem-Solving Methods in Artificial Intelligence* McGraw-Hill. Good coverage of the search and production systems paradigms.

Nils J. Nilsson. 1980. *Principles of Artificial Intelligence.* Tioga Publishing

Marvin Minsky. 1986. *Society of Mind.* New York: Simon and Schuster. A lot of interesting ideas, particularly with respect to parallel distributed systems.

Herbert A. Simon. 1979. *Models of Thought.* New Haven: Yale University Press.

Herbert A. Simon. 1982. *Models of Bounded Rationality.* Cambridge: MIT Press. collected works.

Terry Winograd. 1983. *Language as a Cognitive Process.* Reading, MA: Addison-Wesley. Good introduction to the concepts and approaches used for natural language processing

Patrick Henry Winston. 1984. *Artificial Intelligence.* 2nd edition, Reading, Addison-Wesley. Relatively nontechnical, emphasizing concepts rather than code; widely used as a textbook; may be in a 3rd edition now

Roger C. Schank and Christopher K. Riesback. 1981. *Inside Computer Understanding.* Hillsdale, NJ: Erlebaum Associates.

Schildt, Herbert. 1987. *Artificial Intelligence Using C.* Berkeley: Osborne/McGraw-Hill. Godd source for a lot of basic algorithms.

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

Expert Systems

There is now a very large literature on this, including a number of textbooks and lots of examples of commercial applications. These are just a few books and some articles using expert systems in political science

Robert Axelrod. 1976. *Structure of Decision.* (Princeton: Princeton University Press) ("cognitive mapping" rather than AI per se but a good example of a formal model of political decision-making incorporating how experts organize information about the world)

G. Matthew Bonham and Michael J. Shapiro. 1976. "Explanation of the Unexpected: the Syrian Intervention in Jordan in 1970" pp. 113-141 in Robert Axelrod (ed.) 1976, *The Structure of Decision.* Princeton: Princeton University Press. This did not start out as an expert system but ended up being one—it was an inadvertent test of a foreign policy expert system.

Jaime G. Carbonell. 1978. "POLITICS: Automated Ideological Reasoning." *Cognitive Science* 2:27-51.

Coombs, M. J. 1984. *Development in Expert Systems.* Orlando, FL: Academic Press. From a special issue of International Journal of Man-Machine Studies; lots of examples

Randall Davis and Douglas Lenat. 1982. *Knowledge-Based Systems in Artificial Intelligence.* McGraw-Hill.

Frederick Hayes-Roth, Donald Waterman, and Douglas Lenat. 1984. *Building Expert Systems.* Reading, MA: Addison-Wesley. This is oriented towards building very large systems but has a lot on the basic concepts

Donald Mitchie (ed.) 1982. *Introductory Readings in Expert Systems.* London: Gordon and Breach.

Judea Pearl, 1984. *Heuristics: Intelligent Search Strategies for Computer Problem Solving.* Reading, MA: Addison-Wesley.

Judea Pearl, 1988. *Probabilistic Reasoning in Intelligent Systems.* Palo Alto: Morgan Kaufmann. Both Pearl books are dealing with the problem of decision-making in the presence of noise, which AI tends to deal with quite differently than do statistical approaches

Sawyer, Brian and Dennis Foster. 1986. *Programming Expert Systems in Pascal.* New York: Wiley. All the code you need to construct a basic system using backward chaining as a reasoning mechanism

Tanaka, Akihiko. 1984. *China, China Watching and CHINA-WATCHER.* In Donald A. Sylvan and Steve Chan, Foreign Policy Decision Making: Perception, Cognition and Artificial Intelligence. New York: Praeger, pp 310-344. Rule-based simulation of China

Machine Learning

Forsyth, Richard and Roy Rada. 1986. *Machine Learning: Applications in Expert Systems and Information Retrieval.* New York: Wiley/Halstead.

John H. Holland, Keith J. Holyoak, Richard E. Nisbett and Paul R. Thagard. 1986. *Induction: Processes of Inference, Learning and Discovery.* MIT Press.

Michalski, Ryszaid S., Jamie G. Carbonell and Tom M. Mitchell. 1983. *Machine Learning: An Artificial Intelligence Approach.* Palo Alto: Tioga Publishing.

Schrodt, Philip A. 1987. "Classification of Interstate Conflict Outcomes using a Bootstrapped CLS Algorithm." *International Studies Association*, Washington, April 1987.

Genetic Algorithms

Lawrence Davis. (ed.) 1987. *Genetic Algorithms and Simulated Annealing.* Los Altos, CA: Morgan Kaufmann. Good collection of early GA work; includes Axelrod's experiments on generating the tit-for-tat strategy in IPD using GA's

David S. Goldberg and Amanda L. Thomas. 1986. "Genetic Algorithms: A Bibliography 1962-1986." University, Alabama: Clearinghouse for Genetic Algorithms Report 86001. I've got a copy of this if you want it; its a bit dated now.

John J. Grefenstette. (ed.) 1987. *Genetic Algorithms and their Applications.* (Proceedings of the Second International Conference on Genetic Algorithms) Hillsdale, NJ: Lawrence Erlbaum Associates.

John H. Holland, 1975. *Adaptation in Natural and Artificial Systems.* Ann Arbor: University of Michigan Press. Original Holland book, develops some fairly robust techniques for gene-like learning systems which form the basis for the classifier

Holland, John H. 1986. *Escaping Brittleness: The Possibilities of General Purpose Algorithms Applied to Parallel Rule-Based Systems.* In R.S. Michelski, J.G. Carbonell and T.M. Mitchell (eds.) Machine Learning 2 . Los Altos, CA: Morgan Kaufman. Chapter 20.

Schrodt, Philip A. 1986. "Predicting International Events". *Byte.* 11,11 (November, 1986)

Schrodt, Philip A.1988b. "Short-term Prediction of International Events using a Holland Classifier." *Mathematical Modeling* Summer, 1988.

Neural Networks

David H. Ackley, Geoffrey E. Hinton and Terrance J. Sejnowski. 1985. "A Learning Algorithm for Boltzmann Machines." *Cognitive Science.* 9 pp. 147-169. Presents the algorithm used in the later Hinton-Sejnowski work

Maureen Caudill. 1987. "Neural Networks Primer" *AI Expert* (December, 1987)

Hinton, Geoffrey E. 1985. "Learning in Parallel Networks," *Byte* 9 (April 1) pp. 265-273.

T. Kohonen, 1984. *Self-Organization and Associative Memory.* New York: Springer-Verlag.

David E. Rumelhart and David Zipser, 1985. "Feature Discovery by Competitive Learning." *Cognitive Science* 9,1:75-1112. (excellent literature review with respect to perceptrons and connection machines)

David E. Rumelhart and J. McClelland. 1986. *Parallel Distributed Processing* (two vols.) (MIT Press). Good technical introduction and quite comprehensive, though with a biological emphasis.

Production Systems

Any of the general AI books discussed above will cover this topic in detail. Cited below are a few applications from political science.

Banerjee, Sanjoy. 1986. "The Reproduction of Social Structures: An Artificial Intelligence Model." *Journal of Conflict Resolution* 30,2:22-252.

Hudson, Valerie M. 1987. "Using a Rule-Based Production System to Estimate Foreign Policy Behavior" in Cimbala, Stephen. *Artificial Intelligence and National Security.* Lexington, MA: Lexington Books, pp. 109-131.

Job, Brian L, Douglas Johnson and Eric Selbin. 1987. "A Multi-Agent, Script-based Model of U.S. Foreign Policy Towards Central America." American Political Science Association, Chicago.

David Klahr. 1986. *Production System Models of Learning and Development.* MIT Press.

Stuart Thorson and Donald A. Sylvan, 1982. "Counterfactuals and the Cuban Missile Crisis." *International Studies Quarterly* 26:537-71.

PROLOG

W.F. Clocksin and C.S. Mellish, 1981. *Programming in Prolog.* New York Springer-Verlag. (the original reference on Prolog; it might have defined the language but didn't—there are now many variants)

D.E. Cortesi, 1985. "A Tour of Prolog." *Dr. Dobb's Journal* (March , 1985). General survey of Prolog.

Natural Language

Duffy, Gavan and John C. Mallery. 1986. "RELATUS: An Artificial Intelligence Tool for Natural Language Modeling". International Studies Association, Anaheim. Description of a complex natural language recognition and knowledge representation system.

Brian Hayes, 1985. "A Mechanic's Guide to Grammar." *Computer Language*, vol.2, no's 10-12.

Kimbrell, Roy E. 1985. "English Recognition," *Byte*, Dec., pg. 129. Very detailed guide to language processing.

Schrodt, Philip A. and David Leibsohn. 1985. "An Algorithm for the Classification of WEIS Event Code from WEIS Textual Descriptions." Paper presented at the International Studies Association, Washington. Application of simple keyword techniques to code international events.

Syntactic Pattern Recognition and Sequence Analysis

Hayward J. Alker, James Bennet and Dwain Mefford. 1980. "Generalized Precedent Logics for Resolving Security Dilemmas." *International Interactions* 7;165-200.

Hayward J. Alker and F. Sherman. 1982. "Collective Security-Seeking Practice Since 1945" pp.113-45 in D. Frei (ed.) *Managing International Crises.* Beverly Hills: Sage.

Bennett, Scott and Philip A. Schrodt. 1987. "Grammatical Sequences in COPDAB Events Data." Paper presented at the American Political Science Association, Chicago.

James C. Bezdek, 1981. *Pattern Recognition With Fuzzy Objective Function Algorithms.* New York: Plenum.

K.S. Fu, 1974. *Syntactic Methods in Pattern Recognition.* New York:Academic Press.

K.S. Fu 1982. *Syntactic Pattern Recognition and Applications.* NY:Prentice-Hall. Fu more or less defined the field of syntactic PR in the 1970's and early 1980's: PR has fragmented considerably since then but there are some useful ideas here

Rafael Gonzalez and Michael G. Thomason, 1978. *Syntactic Pattern Recognition: An Introduction.* Reading, MA: Addison-Wesley. (advanced textbook; numerous nonlinguistic examples using Chomsky grammars)

David Sankoff and Joseph B. Kruskal (eds.) 1983. *Time Warps, String Edits and Macromolecules: The Theory and Pratcice of Sequence Comparison.* New York: Addison-Wesley.

Schrodt, Philip A. 1984. "Artificial Intelligence and International Crisis: An Application of Pattern Recognition." Paper presented at the American Political Science Association, Washington, DC , August. Application of Levenschtein metrics to international events data.

Source Code and Programs

Since AI methods are not available in SPSS, and you can't even do them in LISREL (gasp!), one is dependent on either source code or specialized programs. I can offer (via mail or Bitnet) Turbo Pascal source code for the following methods in fairly well-documented form:

Holland classifier. HCLASS, originally written for BYTE, well-documented and user friendly

Genetic algorithms. for solving mixed-strategy zero sum games; a modification of HCLASS; easily modified to handle the IPD problem

Simple neural network. Widrow-Hoff algorithm; currently set up as a spatial pattern recognizer; thoroughly documented

CLS algorithm. embedded in a bootstrapping technique: CLS is the core algorithm for ID3, which is used in most commercial "rule learning" programs; not very well documented

Assorted pattern recognizers used with international events data

In addition, some of the books cited above provide source code in procedural languages (Pascal or C) for other key algorithms:

Sawyer and Foster (1986) gives code for constructing a backward chaining system; Schildt (1987) gives code for a simple AQ algorithm.

I'm not a fan of either LISP or Prolog (nor, increasingly, is the entire commercial AI industry... virtually any AI algorithm can be at least as efficiently implemented in a recursive procedural language [e.g. Pascal or C] as in declarative languages [e.g. ,LISP or Prolog]) but both languages are available relatively inexpensively for experimental purposes. There is a public domain LISP called XLISP which is quite good; Borland International's "Turbo Prolog" is probably the cheapest way to get a serious Prolog though their implementation is significantly different in several ways from Clockstein and Mellish "standard" Prolog.

The National Collegiate Software Clearinghouse (David Garson's outfit at North Carolina State University) has a variety of public-domain expert systems software available. Some of these are simplified versions of commercial systems and are quite sufficient for classroom use. TPM