

THE AGENDA FOR "SOCIAL SCIENCE HISTORY"*

J. Morgan Kousser

I want to take as my texts today statements made to me in correspondence and conversation by two senior quantitative historians. Each statement illustrates what I believe to be misjudgments about the proper methodological priorities for quantitative historians in America today. To spare these historians from publicity which their casual statements were not intended to invite, but mostly to protect myself against reprisal, I shall not name them here.

The first statement arose because I assigned a particular book in my American Political History course. Some of my colleagues, students, and I were critical of the methodology employed in the book, and a student suggested we might reanalyze the data, employing different techniques. The data set, however, was rather obscure and was apparently not available at any major archive. When I wrote to the author, rather brashly asking for a copy of his computer tapes, I was informed that he had "lost interest" in the project after the first year or so and discarded most of the tapes and IBM cards.

I was astonished. Here was a project based on data painstakingly collected for more than a decade from a wide variety of published and unpublished sources; coded, punched, and verified; and he had chucked it out almost immediately after performing a rather idiosyncratic analysis which had produced striking and controversial conclusions. However obscure the sources, the data could easily have been stored on one inexpensive computer tape. But since he had erased the tape, his analysis could not be checked, much less improved upon, without investing years of drudgery reassembling the data.

Before examining further this first case, let me sketch a second. While talking with another senior historian at a professional meeting, I was asked what I had been doing recently. I told him I had been dissatisfied with the rather murky understanding of statistics I had gotten in graduate school, and that I had therefore been retooling by taking mathematics, statistics, and econometrics courses. His response was to ask me what good it did to know calculus and matrix algebra. Again, innocent that I am, my response was astonishment. I had thought that everyone gave at least lip service to the ideas that, first,

* This paper was first given to a panel on "Priorities in American Behavioral History" at the SSHA meeting in Madison, Wisconsin, April 23-24, 1976. I have retained its rhetorical and slightly hyperbolic character deliberately in order to provoke controversy.

math was a good thing; and, second, it was better to understand techniques deeply rather than superficially. Unprepared for his query, I responded only lamely and vaguely.

I want to use these two minor events as jumping-off points for making a few concrete and, I hope, practical suggestions about the proper methodological agenda for historians in the next few years. I will contend that: (1) an excess of individualism has led to a squandering of opportunities for accumulating and improving both our knowledge of the past and the profession's overall level of methodological expertise; (2) funding agencies and consulting committees for those agencies could greatly assist the profession by further encouraging the storage of data collected under grants from them and the initiation of projects based on reanalyses of already available data; (3) by applying now some lessons which may be gleaned from the experiences of our sister social science disciplines, we may avoid merely recapitulating what are now seen in those disciplines as unfortunate side-tracks and errors; (4) except for some projects in economic history, the techniques employed by self-proclaimed quantitative historians have been a good deal cruder than they could and ought to have been; (5) the use of more sophisticated techniques might well have led to different conclusions; and (6) the exploitation of these techniques will require a radical change in the way quantitative historians are trained and, probably, will also require the development of methodology into a full-blown subdiscipline in the profession.

Let me argue the first three points by reanalyzing the experience of the state and local politics subdivision of political science in the 1960s.¹

In 1960, state and local politics was a backwater which had resisted the behavioral flood of the 1950s. Members of this land-locked subdiscipline rejected the tendency of the behavioralists to generalize broadly, just as they hesitated to immerse themselves in the deep waters of sociological and psychological theory and the tide of statistics that characterized the new reign in the profession. In 1963, however, James A. Robinson and Richard E. Dawson plunged into the behavioral waters to rescue a bobbing remark which V. O. Key had cast forth some fourteen years earlier. Assembling from data in the *Book of the States* a set of variables with which they sought to operationalize and test Key's casually-formulated hypothesis, they launched a whole literature on "policy outputs"—a literature which is still sailing along.

Undergirding this literature is a framework of factual and theoretical assumptions which were never very clearly blueprinted, and this failure to formulate clearly and examine closely the underlying design of the work made the whole project extremely unseaworthy from the beginning. Let me offer one possible formulation of the underlying

axioms. As I see it, there were six crucial postulates which made up the empirical theory behind most of the state and local policy outputs literature: (1) sets of elites who care about little but winning complete in the political system by bidding for the votes of groups of non-elites; (2) the vast majority of non-elites rationally decide which package of political goods to purchase by choosing that bundle which maximizes their economic welfare, narrowly defined; (3) the system operates so as to offer a premium to those who declare their choices last; (4) members of the lower socioeconomic classes are less likely to vote than those in higher status groups; (5) virtually all governmental expenditures are redistributive: the higher the government budget level, the better off the lower classes are, compared to the upper groups; and (6) nearly everyone has sufficient information about the effect of government spending to make their voting calculations.

Although parts of these assumptions were themselves testable, most of the literature concentrated on hypotheses derived from them: increasing party competition should have drawn more lower class voters into the political system. Greater turnout would therefore reflect greater participation by the relatively poor. Since the lower strata would be the last to decide, the price of their votes would be maximal. Because governmental expenditures were redistributive, the lower classes would demand higher budgets in payment for their franchises. Therefore, higher degrees of party competition and turnout should be associated with higher levels of government expenditures. The null hypothesis could be formulated in two ways: either governmental expenditures of all kinds had roughly the same weight in everyone's utility function, or they had the same weight in the utility functions of the elites in every geographical area and those elites, especially bureaucrats, paid little attention to what mere voters wanted.

The data set on which these competing theories were tested consisted essentially of appropriately normalized indices of governmental expenditures, income, party competition, turnout, education, etc., for each state in the post-World War II United States. At first the analyses were methodologically quite simple: states were simply classified into subgroups on various indices and those groupings were set side by side to see how they fit. This technique shortly gave way to rank-order and Pearsonian correlation coefficients, to multiple and partial correlation, to multiple linear regression, stagewise regression, factor analysis, and nonlinear regression. The essential data set was quickly passed around to virtually all major universities and was constantly extended by adding more variables and extending the time series. A great many graduate students and young professors appear to have cut their statistical teeth on this field; and for a while it became *de rigeur* for each issue of a political science journal to contain an article on policy outputs which

either introduced a previously untried statistical technique, extended the data set, refined an index, or critiqued a previous article.

I would like to make two points about the policy outputs literature. First, the reanalyses did train a good many students and increase the general level of methodological competence among political scientists. This rise in competence made more sophisticated tests of theories possible. It did not, however, add as much as it could have to the profession's understanding of the policy process, because most of the contributors to the literature were content merely to crunch data. Since they did not think very seriously or deeply about their theoretical assumptions, they did not get very far in their theoretical conclusions. They did not, for example, explicitly mesh their research with the emerging formal theories of electoral competition and turnout derived from the work of Anthony Downs.² The second point, then, is that the full benefits of reanalyses—or initial analyses, for that matter—will be achieved only when theory and methodology proceed hand in hand.

To summarize so far: the lesson I draw from the policy outputs literature is that reanalyses of widely available data sets offer major opportunities for the advancement of a profession, but the benefits will be directly proportional to the theoretical self-consciousness of the effort as a whole.

Let me turn now from reanalyzing reanalyses to considering past efforts in quantitative history and proposing changes in the way methodology is taught in history, and to considering the weight which should be attached to technical competence.

In a recent article in *Reviews in American History*, I examined the methods employed in three leading works in quantitative history, especially efforts by members of the so-called "ethnocultural school."³ Let me repeat myself on only a couple of points. I found that, at a time when canned multiple regression programs were widely available, when both exposés of the so-called "ecological correlation fallacy" and methods for dealing with it were pretty widely known, Professors Formisano, Hackney, and Kleppner relied in their monographs chiefly on so-called "eyeball correlation," and (when they did employ statistical techniques) on rank-order coefficients, and bivariate Pearsonian correlation coefficients. Perhaps more importantly, they generally assumed that the relationships of interest were all nicely linear and correctly specified; that in effect, limitations on the range of their independent and dependent variables affected neither the bias nor the efficiency of their parameter estimates. In other words, not only was the methodology employed much less sophisticated than it could easily have been, but better techniques almost certainly would have led to different or at least much weaker conclusions.

Let me explain the point about truncated or limited dependent variables more fully and use it to illustrate my argument that it is necessary to include a much larger component of mathematics in training quantitative historians.

In their analyses of the effects of the ethnic and religious composition of the population on voting behavior, many of the early ethnoculturalist historians focused on "banner units"—e.g., the most Democratic or most Whig precincts or townships in an area. By concentrating on these units, they were considering only a limited range of their dependent variables. It is easy to show that this leads to a bias in the parameter estimates when one uses ordinary least squares or some simpler method which can be treated as an example of OLS. Thus, because they employed a faulty methodology, the conclusions of some leading quantitative historians are at best uncertain and at worst incorrect.

Now, there has been a good deal of work on the problem of limited dependent variables in biometrics and, recently, in econometrics. The proposed cures for the disease of limited observations on the dependent variable involve probit, logit, Tobit, and other forms of *-it* analysis. The trouble is that such techniques often require the use of iterative algorithms to solve analytically insoluble integrals in equations which maximize various likelihood functions. Such forms of analysis—which, by the way, have considerably wider potential applications in political and social history than I have time to outline here—will be pretty incomprehensible to historians who lack a solid grounding in calculus and statistics. Yet since simple techniques, mechanically applied, will often in practice lead to misleading or incorrect results, a much deeper knowledge of mathematics and statistics becomes a prerequisite both for *doing* good quantitative history and for evaluating the findings of quantitative studies. Even historians who do little quantitative work themselves, in other words, need a much firmer grounding in methodology to be able to understand and appraise the works of other social, political, and especially economic historians. How can historians continue to stomach reviews of quantitative works which begin with some version of the disclaimer: "I don't know anything about numbers but I know what I do (or don't) like"!⁴

Recent experience from other social science disciplines gives even greater emphasis to the professional disutility of a mere formulaic and superficial statistical learning. Psychology and sociology departments, which have long required simple statistics courses of their graduate students, increasingly seem to be requiring calculus as an entry prerequisite and using it in their statistical sequence. Political scientists, moreover, are growingly critical of the sixties' tendency to toss lots of

data into a computer, add a dash of canned program, and serve up the resulting indigestible concoction to the public. The "workshop" articles in the *American Journal of Political Science* and the increasing tendency to substitute Johnston's *Econometric Methods* for Key's *Primer of Statistics*, Blalock's *Social Statistics*, or similar works evidence the fact that, in political science at least, a cookbook knowledge of statistics won't get you into heaven or even into a job any more.⁵

To profit from the past experience of other disciplines, historians should move as quickly as possible through the cookbook phase to the stage in which self-conscious theorizing guides sophisticated methodological tools whose use is well understood. Learning from the example of economic history, which has almost everywhere except at the University of Wisconsin been amputated from history departments and grafted onto economics, we should also attempt to avoid the further dismemberment of history as a discipline. Quantitative economic history is, as the brouhaha over *Time on the Cross* demonstrates, too important to the historical profession to be left to a small group of cliometricians. And though there have as yet been no bombshells in quantitative political or social history which have left as large craters as *Time on the Cross* has, the same argument holds for those subdisciplines as well. In sum, unless history departments begin to make a sophisticated training in statistical methodology and the relevant economic, political, or social theory a central part of their graduate training, they will be perpetuating easily overcome errors and producing students who will be increasingly unable to comprehend major developments in their fields.

Such a program for the future of the historical guild will, of course, require a more mathematical apprenticeship and a larger number of well-trained masters than we have in the profession now. How are we to get from here to there? I have several specific proposals:

1. History graduate programs should encourage applications from undergraduates with mathematical training. (There do exist some math and even engineering undergraduates who are not totally illiterate.)

2. Temporarily, until they can offer the relevant training themselves, history departments should relax requirements in graduate programs to encourage students to tool up in mathematics and to take theory and methods courses in economics, political science, sociology, psychology, or statistics departments. Social and political historians especially ought to be urged, as their colleagues in economic history are now, to get a firm grounding in microeconomic theory and some of its applications. Microtheory should serve not only as a source of hypotheses, but more importantly as a paradigm of how explicit theorizing and model-building might proceed. Students who wish to be intensively trained in new areas should neither be forced to take course overloads

nor to pick up everything piecemeal and on their own. The requirement that all graduate students take minors in fields far outside their areas of primary interest must obviously fall. It may be cute for an American social historian to be able to discuss T'ang poetry or to name all the popes from the eighth through the fourteenth centuries, but these "broadening" facts are certainly less essential to his work than knowing what a partial derivative or a biased estimator is.

Acquiring the relevant tools in statistics and theory ought to be viewed as complementary rather than contradictory endeavors. For one thing, the same mathematical background is useful to both. For another, the formulation and testing of theory, or at least testing the consequences of theory, are logically very strongly related. One obviously cannot operationalize and test theoretically derived hypotheses without knowing something about theory; conversely, untestable or untested theories will not interest historians or other empirical social scientists very much. Theory and methodology are like the Emperor and his new clothes; without the one, you have an empty suit; without the other, a nude old man:

3. Senior historians should encourage foundations, government agencies, and major grantees to set up postdoctoral programs to allow historians to deepen their knowledge of methods and theory, or to broaden existing programs to include historians.

4. Presently-working historians need to realize that summers at the Newberry or Ann Arbor, however useful in curing their irrational phobias about numbers and computers and introducing a few techniques, can no more transmogrify statistical illiterates into expert data analysts than a Berlitz course of the same length can metamorphose an American into a Russian historian. Those who expect to be able to do more than read a simple menu in the language of statistics must consume calculus and linear algebra as appetizers and polish off a healthy main course of statistics and econometrics so that they can choose a dessert from a list of special applied topics such as probit and logit, scaling, or simultaneous equations. Short cuts or snacks will only result in indigestion or methodological anemia.

5. History departments, journal editors, and referees must comprehend both the difficulty of methodological retraining and the importance to the discipline of methodological innovation. Since it will take at least three years of intensive work (while carrying a teaching load) to achieve a passable level of methodological competence, methodologists-in-training may produce relatively few publications. The promise of innovation should therefore be weighed more heavily in their promotion decisions than it often is today. Substantive articles which employ methodologies new to the profession, as well as nonsubstantive articles expositing such techniques, ought to grace the pages

not only of the interdisciplinary but also of the main-line journals. The NEH, the NSF, and the referees and committees for those agencies ought to give special preference to methodologically innovative projects and to require that any new machine-readable data produced in those projects be deposited in compatible and comprehensible formats at some central depository such as the one in Ann Arbor.

6. In the medium run, say, ten years, cliometricians (meaning here a breed of historians well schooled in relevant statistical techniques) ought to compose a small but well recognized subdiscipline in history similar to econometrics and the new polimetrics.

Why a new subdiscipline? Why can't historians with quantitative data sets just go visit their friendly local statisticians or econometricians when they have questions about methodology? Besides my general belief that artisans should fully appreciate the tools they employ, there is also the fact that the problems historians will face will differ somewhat from those faced by other disciplines. Econometrics does not pay much attention to contingency tables or Markov processes; quantitative historians do or will. Polimetrics puts much more emphasis on analyzing surveys, psychometrics on analyzing batteries of tests, sociometrics on analyzing intragroup microrelations than most historians probably will. Moreover, certain topics which are of marginal interest in other fields, such as geographical aggregation problems, will be central in cliometrics. The training of students will be more efficient and advances on topics of special interest to historians will be more probable, therefore, if we begin to build up a specialized subdiscipline of our own. A first step toward the recognition and organization of this subdiscipline has now been taken by setting up a methodology network in the Social Science History Association.

7. In the long term, say, a generation, every social, economic, and political historian should be expected, as a matter of course, to have a firm grasp of statistics and the theory from at least one social science discipline. For that development to yield really interesting, solid results for history to become a science—that is, a discipline firmly grounded in theory, in which results are falsifiable and cumulative—we must grant first priority on the historical profession's agenda to serious, sophisticated training in both theory and methods.

J. Morgan Kousser
Associate Professor of History
California Institute of Technology

NOTES

1 This ongoing literature is much too vast to cite here. For a discussion of recent work in the area, see Herbert Jacob and Kenneth N. Vines, *Politics in The American States*, third ed. (Boston, 1976).

2 See Downs' *An Economic Theory of Democracy* (New York, 1957), and William H. Riker and Peter Ordeshook, *Introduction to Positive Political Theory* (Englewood Cliffs, N.J., 1973).

3 "The 'New Political History': A Methodological Critique," *Reviews in American History*, IV (1976), 1-14.

4 Examples might be drawn from reviews of Robert W. Fogel and Stanley L. Engerman, *Time on the Cross* (Boston, 1974), or of my own book, *The Shaping of Southern Politics* (New Haven, Conn., 1974). Again, in order to minimize the number of my enemies, I shall not cite specific examples.

5 John Johnston, *Econometric Methods*, second ed. (New York, 1972); V.O. Key, Jr., *Primer of Statistics for Political Scientists* (New York, 1966); Hubert M. Blalock, *Social Statistics*, second ed. (New York, 1972). See also the new moderately mathematical text by Eric Hanushek and John Jackson, *Statistical Methods for Social Scientists* (New York, 1976).