Quantification in History

WILLIAM O. AYDELOTTE

OVER the past generation a number of historians have recognized that counting, when circumstances permit it, may assist in the explanation of a limited class of historical problems. The historical monographs in which quantitative methods have been used are already sufficiently numerous so that a review of them would require an article by itself. The purpose here is not to survey this literature but, instead, to raise several general questions related to it. Professional opinion regarding the value of quantification for history has been rather less than unanimous, and discussion of the subject has occasionally been acrimonious. There have also been a few misunderstandings. I wish to consider here what is involved in trying to apply quantitative methods to history, what kinds of results may be expected, and what difficulties lie in the way. Though I shall say something about the advantages of quantification, I am also, in a sense, concerned to speak against it and to make clear the problems it presents. My own approach to the subject is

* A professor at the University of Iowa, Mr. Aydelotte is interested primarily in modern British history. He has written, among other things, “Voting Patterns in the British House of Commons in the 1840s,” Comparative Studies in Society and History, V (Jan. 1963).
Some historians, of course, disclaim any intention of making such state-
ations conservative and skeptical, and at times I feel that the current fad for quan-
tification has been pushed too far. In any case, the exploration of the limita-
tions of a method is an effective device for revealing its characteristic features.
The principal value of quantification for the study of history, stated in the simplest terms, is that it provides a means of verifying general statements. Some historians, of course, disclaim any intention of making such state-
ments and insist that the business of a historian is not to generalize but to tell a story. Such a view can hardly be seriously entertained as a description of the objectives of all historians, for it manifestly does not apply to the work of a number of eminent members of the profession. One might question, indeed, whether any historian can avoid generalizing altogether. It is an idle task, however, to attempt a formal prescription of a historian's duties. If some wish to emphasize narrative more than others, there is no reason why they should not. History is what historians do, and they do different things. It would be presumptuous to dismiss any of their objectives as being in some fashion improper. The day of a single methodology in history, if it ever existed, is at any rate now gone. In a discipline where there are at present so much upheaval, reassessment of methods and values, and introduction of new approaches, it seems better to say that anything historians do is useful if it can be shown to be useful.

For historians who do wish to generalize, however, quantitative methods can offer certain advantages. Generalizations are implicitly quantitative in character, even though this may not always be clearly brought out. As Lee Benson says, historians who use words like “typical,” “representative,” “sig-
ificant,” “widespread,” “growing,” or “intense” are making quantitative statements whether or not they present figures to justify their assertions. Unfortunately, not all historians seem to realize the need to check general statements. Benson complains, in the same passage, of “the impressionistic approach long dominant in American historiography,” and I have oc-
casionally been bothered by this kind of thing in my own field. Historians justly pride themselves on their techniques of verification, which have be-
come in some areas highly sophisticated. It seems fair to say, however, that these techniques have more often been applied to individual bits of in-
formation than to broader statements. Some writers, after a precise descrip-

1 The eleven contributors to a recent volume of essays on this subject, as the editor states in his summary: “all agree that the historian willy-nilly uses generalizations at different levels and of different kinds.” (Generalization in the Writing of History, ed. Louis Gottschalk [Chicago, 1961], 208; see also, on this point, Alfred Cobban, The Social Interpretation of the French Revolution [Cambridge, Eng., 1964], 5–7.)

tion of a few cases, will proceed to generalize blithely about the motives of large groups of men even though the evidence to support their views is often not presented and, indeed, would be hard to come by, for the motives of most men are obscure and not easy to discern. To an uncritical audience several concrete illustrations may carry more conviction than a statistical table. Yet to support an argument by only a few examples, though it may be a persuasive rhetorical device, is not logically adequate. There are exceptions to most historical generalizations, and, if the citation of occasional instances were accepted as proof, it would be possible to prove almost anything.

Quantitative methods, the numerical summary of comparable data, make it possible, in some cases, to avoid these pitfalls. The condensation of data by such means, when it is clearly legitimate, constitutes a saving of time and a convenience in that it makes the information easier to describe and to handle. It also helps to ensure a greater degree of accuracy. Memory is selective, and general impressions are notoriously untrustworthy. When the data are so numerous that they cannot all be kept clearly in mind at once, the investigator is likely to remember best the cases that fit his own preconceptions or his pet hypotheses. An orderly presentation of the evidence in quantitative form helps the student to escape the tricks that his memory plays upon him. Quantitative analyses are, of course, gratuitous when the number of cases is small, when the student is concerned with only a few men or, perhaps, one man, and when the general tenor of the materials can be immediately grasped. As the data become more numerous, however, a systematic arrangement of them becomes the more desirable. There are, indeed, some questions, of which examples will be given presently, which could hardly be attacked without the use of methods of this kind.

A quantitative presentation of the available information can help to direct the student's attention to the questions most worth investigating. Since it brings the whole of the evidence, on the point it covers, into intelligible focus, the general character of the findings can be more readily perceived and relationships and differences emerge that could not so easily have been observed without this reduction of the data. Such an analysis reveals what events or issues were of special interest, in the sense of involving change through time or departure from the norm, and hence might particularly repay investigation. It can, in this manner, help in defining or restating the historical problem to be studied.

Beyond this, a quantitative analysis offers a systematic means of testing hypotheses. It establishes how many examples there are to support each side
of the argument and thus reveals not only the main features of the evidence but also, more important, the exceptions to them, the nuances, the degree to which the emerging generalizations need to be qualified. Measurement locates the defect in the original hypothesis and registers "the departure from theory with an authority and finesse that no qualitative technique can duplicate." A quantitative discrepancy between theory and observation is obtrusive. "No crisis is . . . so hard to suppress as one that derives from a quantitative anomaly that has resisted all the usual efforts at reconciliation."8

The general overview of the whole evidence obtained by quantitative means can also be a powerful stimulus toward the reformulation of one's ideas. When anomalies occur, the student can direct his attention to the cases that do not fit the original theory, try to find out why they are exceptional, and, by rearrangements of the data, test alternative hypotheses that may account for a larger proportion of the evidence. Such manipulations of the data would take an immense amount of time to do by hand, but, ordinarily, they can readily be performed by machines. I advise my students, if they are working with fifty cases or more, to punch the information. This is easily done, and, once it is done, there is no great difficulty about trying additional correlations. By the same token a quantitative analysis can even, in some cases, point the way to the formulation of new hypotheses that will make the findings more intelligible.

The case for quantification might be made in still a different way by saying that it is a method of reasoning, one that involves number. As one of my colleagues at the University of Iowa has put it, quantification adds, to whatever factual or historical premises may have been established, the premises of mathematics as well. "Arithmetic is a vast treasure house of additional premises, or, what amounts to the same thing, of patterns of deductive inference. Quantification is the key to the treasure."4

The advantages of this approach have been appreciated by a number of present-day historians. G. Kitson Clark suggests as appropriate advice to someone who wishes to generalize about a group or a class: "do not guess, try to count, and if you cannot count admit that you are guessing."5 Lawrence Stone writes: "Owing to the obstinate perversity of human nature, it would no doubt be possible in England of 1958 to find, if one tried, declining manual labourers and rising landed gentry. To have any validity at all, conclusions about social movements must have a statistical basis."6

---

9 Gustav Bergmann, Philosophy of Science (Madison, Wis., 1957), 69.
11 Lawrence Stone, letter to editor, Encounter, XI (July 1958), 73.
Applications of quantitative techniques to historical materials have, in some cases, materially advanced the discussion of major problems. Monographs on the composition of the British House of Commons, which are now fairly numerous and cover a span of six centuries, have brought to light significant continuities and changes in the social structure of the British political elite. Crane Brinton, in his well-known quantitative study of the members of the Jacobin Clubs, reached the conclusion that the Jacobins represented "a complete cross-section of their community" and that: "The Jacobins of 1794 were not a class, and their enemies the 'aristocrats' were not a class; the Terror was not chiefly then a phase of the class-struggle, but even more a civil war, a religious war." Donald Greer, on the basis of a quantitative analysis of the victims of the Terror, argued that the lower classes, by the definitions he used, supplied 70 per cent of the victims and the upper classes less than 30 per cent and that: "The split in society was perpendicular, not horizontal. The Terror was an intra-class, not an interclass, war." From the researches of Brinton, Greer, and others, crude class theories about the French Revolution have received a setback. Revisions have also been made in accepted views about American history. Richard P. McCormick published in the *American Historical Review* a set of tables, drawn from readily available election statistics, on the basis of which he was able to show that the great popular turnout of 1824 was a myth and that: "In the 1824 election not a single one of the eighteen states in which the electors were chosen by popular vote attained the percentage of voter participation that had been reached before 1824." His finding contradicts the assertion he quotes from a standard text that, in the period before 1824, "only small numbers of citizens seem to have bothered to go to the polls." It contrasts also with Charles and Mary Beard's colorful statement that, by 1824, "the roaring flood of the new democracy was now foaming perilously near the crest ..." and with Arthur M. Schlesinger, Jr.'s reference to the "immense popular vote" received by Jackson in 1824. Albert Ludwig Kohlmeier, using statistical data on canal and riverboat traffic, was able to show when and how rapidly the trade of the Old Northwest shifted away from the South and to the Northeast. Stephan Thernstrom, by a quantitative analysis based largely

---

10 Albert Ludwig Kohlmeier, *The Old North-West as the Keystone of the Arch of American Federal Union: A Study in Commerce and Politics* (Bloomington, Ind., 1938).
on census records, exploded various familiar hypotheses about social mobility in a Massachusetts town in the later nineteenth century. Quantitative presentations have formed the basis for substantial generalizations by an impressive group of additional historians including Thomas B. Alexander, Bernard and Lotte Bailyn, Allan G. Bogue, Jean Delumeau, Robert W. Fogel, Frank L. Owsley, Lawrence Stone, Charles Tilly, Sylvia L. Thrupp, and Sam B. Warner, Jr. This list of examples could be considerably extended.

These results have often been achieved by fairly simple methods; for much historical research the quantitative procedures required are not complex. Historians do not ordinarily need to deal with problems of statistical inference in which an attempt is made to ascertain the characteristics of a large population by inspection of relatively small samples. Their work is usually limited to the easier task of descriptive statistics in which the object is to portray the characteristics of a group, all members of which have been studied, and to correlate some of these characteristics with each other. The computations needed for this are not ambitious. All that is generally required are a few totals, a few percentages, and a few correlations in which the relationship between certain variables is examined while other variables are controlled. This is a simple matter mathematically, although the research may be laborious, and it is simple mechanically as well. Even so modest a use of quantitative methods can sometimes produce results of great interest and can be used to test historical generalizations of some scope on which there has heretofore been scholarly disagreement. Since only a limited amount of such research has been done, much gold is still near the surface. It may turn out, however, that richer veins lie deeper. Though it has proved extremely useful to classify, arrange, and summarize the available information, it may be even more rewarding—to judge from some of the ventures that

have already been made—to attempt more complex methods of descriptive statistical analysis by the use, for example, of mathematical models or of scaling techniques. 13

Although substantial and interesting work has been done along these lines, much more could be attempted. Historians who have used quantitative methods have been timid in their application of them and have come nowhere near exploiting their full potentialities. Also, many historians who deal with problems for which such methods might be helpful have not tried to use them at all. Economic history is, perhaps, an exception. This field is naturally suited to quantitative research since many of the original data come in quantified form, the problems and hypotheses tend to assume a quantified shape, and, in the field of economics, theoretical analysis is more advanced. In political and social history, however, opportunities have been missed. Though the area of historical research to which these methods can be applied may be limited, it has certainly not yet been fully explored.

Furthermore, much hostility to quantitative methods still remains among some members of the historical profession. Despite what might seem the obvious advantages of these methods for certain kinds of problems, despite their notably successful application in many historical projects, and despite their long acceptance as a matter of course in several related disciplines, some historians still object to them vociferously and consider them altogether inappropriate for historical research. Questions have been raised regarding: (1) the value of the work that has been done; (2) the feasibility of this approach in view of the admittedly limited materials available to historians; (3) the reliability of the results obtained by these techniques; and (4) the usefulness or significance of the results. These objections are not wholly without foundation. It would be pointless to deny either the limitations of the method or the lapses of some of its practitioners. To concede this, however, is not to tell the whole story.

(1) Certainly the ventures of historians into this kind of research have not been uniformly fortunate. Some of these studies, far from revolutionizing historical thought, have themselves not stood the test of time and have been shown to contain imperfections of method, which, to some extent, vitiate

their conclusions. It would be unfair to mention individual monographs without a more extended discussion of their arguments than is possible in this paper. I shall have occasion to describe some of the statistical solecisms committed by a few workers in my own field in separate articles on special topics. In general, it has been contended, sometimes plausibly, that a number of the pioneers in quantitative historical research overlooked certain elementary precautions. They did not, it is said, always appreciate or remember that a sample that is small and, hence, biased or unrepresentative may distort the results, that percentages should be figured in terms of what is hypothesized to be the independent variable, that a conscientious search should be made for all possible relevant variables (though it is unlikely that they can all be found), that failure to make such a search may produce spurious correlations, or that refinements of technique cannot compensate for the inaccuracy or incompleteness of the original data. Doubtless the application of quantitative techniques to history has not paid off as well as might have been expected because of the statistical naiveté of a few of those who first tried it. To say this, however, is not to disparage quantitative methods. On the contrary, these are exactly the errors that an experienced statistician would not commit, and they arise not from an overemphasis but from an underemphasis on accepted statistical procedures. It may not be unreasonable to expect that simple technical errors of this kind will occur less frequently in the future as a new generation of historians becomes more alert to what is needed for this type of work.

(2) A more serious objection is that quantitative techniques may not be feasible at all in history, or can be used only within narrow limits, because of the complexity of historical materials and the restrictions on historical knowledge. It is difficult to get accurate information, for the sources may prove inconsistent or unreliable. Also the task of correctly recording so great a mass of data is more arduous than is likely to be believed by anyone who has not tried it; the natural proclivity of almost all men to error, to incorrect observation, has been repeatedly shown by experiment. Beyond this, however, there are formidable problems of taxonomy. A given body of data can generally be classified in any of many different ways, and skill and experience are needed to choose the categories that will prove most useful. Unfortunately it may not become apparent which these are until one is well into the research and it is too late to change. It is also no easy matter to make the categories precise and clearly distinguishable from one another. The existing vocabulary of social history is inexact, and many of the terms in common usage are too vague to permit unequivocal classification of the data.
To give one example, problems of this kind have, according to a recent review article, bedeviled research on the supposed conflict of the aristocracy and the bourgeoisie during the French Revolution. The ambiguities in the definitions of these terms have had the result that:

the central doctrine of the class struggle between bourgeois and aristocrats can now only be accepted as an act of faith; for no two people can agree on who the bourgeoisie and the aristocrats were; no one can formulate (and few even try to formulate) a criterion for distinguishing between them that can be followed consistently, and every argument is thus liable to be at variance with easily ascertainable facts.14

Similar problems arise, of course, in the social history of other countries. If a historian tries to distribute a group of men among conventional categories of this kind, borderline cases may necessitate so many subjective judgments that the resulting classifications will not be worth much. No amount of expertise in the manipulation of the figures will make adequate correction for imprecision in the original data or for categories that do not adequately measure what it is claimed that they measure. A quantitative approach does not of itself ensure accuracy. Jeremy Bentham’s "felicific calculus" was set forth in quantitative terms, but it is not generally regarded as a precise conceptual scheme. There is a danger, in this kind of work, of a spurious precision—giving the results, to several decimal places, of calculations based on incorrect original assumptions. If the classifications used at the start are worthless, the computations based upon them will be equally so, no matter how many times they are passed through the computer, and the situation will develop which is known in the trade as "GIGO": "garbage in and garbage out."

Furthermore, historical information is restricted. Historians who seek to use quantitative methods are, in comparison to those working with contemporary affairs, at a disadvantage. It is difficult—and the difficulty generally increases with the remoteness of the period studied—to obtain relevant data for a large enough sample of the group or "population" under consideration to make a quantitative presentation useful and effective. It is feasible, for example, to study the composition of the British House of Commons in recent centuries, though the task becomes harder as one goes back in time, but it might be less rewarding to attempt an analysis of the personnel of Justinian’s army.

14 Betty Behrens, "'Straight History' and 'History in Depth': The Experience of Writers on Eighteenth-Century France," *Historical Journal,* VIII (No. 1, 1965), 125; see Greer’s comments on the ambiguities of his own categories, in *Incidence of the Terror,* 88–96; for a more extended discussion of these problems, see Cobban, *Social Interpretation of the French Revolution,* Chaps. iii, vi, viii–xiv.
Even some of the historians who have made conspicuously successful use of these methods complain frequently about the inadequacy of the sources with which they had to work. Brinton found the membership lists of the Jacobin Clubs incomplete, a problem heightened by the considerable turnover in membership, while the occupations of some Jacobins were not listed and the occupations of others were described in ambiguous terms. He insists that part of his information does not “have even the relative accuracy possible in a study of contemporary demography.” Greer speaks of “the impossibility of determining with any degree of exactitude the total death roll of the Terror.” Owsley found that the tax lists for many large areas of the South had not survived, while the census reports, besides being less accurate as a rule than the tax lists, were seriously incomplete except for the latter part of the ante bellum period. Warner found rich statistical materials surviving for nineteenth-century Boston, but noted that city, state, and federal counts did not agree with each other and added the warning that: “The presence of substantial errors in the census requires the local historian to use census data with the same sophistication he would use any other source. The past tendency to check writings of individuals against other sources but to accept statistics as primum facie fact must be abandoned.”

Even in cases of groups for which quantitative methods can to some extent be used, it is not always possible to employ tests of sufficient refinement to verify what appear to be the most significant hypotheses. Benson, in his discussion of Beard’s interpretation of the battle over the Constitution, has suggested that “we are likely to progress further if we group men, not according to their ‘economic interests,’ but according to various other things including, for example, ‘their values, their beliefs, their symbols, their sense of identity.’ Yet it may not be easy to obtain this kind of information for all or most members of a population of any size removed at some distance in time. It could, perhaps, be found for a few individuals on whom detailed information can be gathered from their correspondence and papers, but in statistics arguing from a few not necessarily representative examples is the great heresy. Doubtless more can be done than has always been realized, and, in another book, Benson has applied imagination and ingenuity to available materials, to draw impressive and persuasive inferences on some of
Quantification in History

these difficult matters.\textsuperscript{20} It can scarcely be denied, however, that these are obdurate questions, and anyone who tries to solve them has his work cut out for him; ingenuity can carry only a limited distance. For many groups in the past the kind of information needed to make such tests, much of it at least, has long since disappeared and is now irretrievable.

Clearly, formal statistical presentations are feasible only for a limited range of historical problems. The available information may be insufficient or may contain ambiguities that make it difficult to summarize in intelligible categories. Nothing is to be gained by pretending otherwise or by attempting to force the use of these methods beyond where evidence will carry. Frank Knight once observed that the dictum attributed to Lord Kelvin—"If you cannot measure, your knowledge is meagre and unsatisfactory"—has in practice been translated into the injunction: "If you cannot measure, measure anyhow."\textsuperscript{21} This, of course, would be a counsel of darkness. Whether quantitative methods will be helpful on a given problem is a matter not of rule but of the strategy of research.

Though these difficulties are substantial, it would be a gross distortion to regard them as insuperable. Taxonomic problems vary in incidence, and it is mistaken to suppose that all subjects are equally difficult to quantify. Social categories may be tricky, but other kinds of information, such as votes in a legislative body, can be tabulated with some assurance. Economic and demographic data have been handled quantitatively with success for some time.

Even in the study of social history it has sometimes proved possible, as it has in scientific investigation,\textsuperscript{22} to advance the argument by jettisoning subjective definitions and adopting objective ones, by disregarding earlier concepts that were too vaguely defined to admit of measurement, and by concentrating instead on categories that could be unmistakably specified—not "aristocrats," but peers and their sons; not "gentry," but men included by John Burke in his reference work The Landed Gentry of Great Britain and Ireland; not "businessmen," but men engaged in certain ways in certain types of business. Whether these more sharply defined categories correspond accurately to the old categories is a question that cannot be answered since the old ones are so indefinite that they cannot be said to correspond accu-


\textsuperscript{21}Kuhn, "Measurement," 31, 34; remarks by Frank H. Knight in Eleven Twenty-Six: A Decade of Social Science Research, ed. Louis Wirth (Chicago, 1949), 169. The quotation ascribed to Kelvin appears on the façade of the Social Science Building at the University of Chicago. Kuhn has been unable to find these exact words in Kelvin's writings, though Kelvin expressed the idea more than once in slightly different language.

\textsuperscript{22}See the discussion of the development of the concept "degree of heat" in Kuhn, "Measurement," 58–59.
rately to anything. One cannot, by using the new categories, effectively test propositions couched in terms of the old ones. Such propositions cannot, indeed, be tested at all, for an imprecise or slipshod formulation is impregnable; a statement that has no exact meaning cannot be disproved. What is feasible, however, is to study a group or an entity that might be conjectured to correspond somewhat to the old and loosely defined concept but that at least has the virtue that it can be identified. The investigator must, of course, assume the burden of showing that his new categories are viable and useful. The great step forward is to take the objective or unequivocal definition as the norm, as describing the entity that will be subjected to analysis, and to demote the subjective or vague concept to a subordinate position, to appreciate that, though it may serve as a useful starting point in the formulation of an operational definition, it may also contain variables that are difficult to measure or even to identify and that it cannot, therefore, be handled in any conclusive fashion. By this procedure one at least knows where one stands, and the problems of social measurement may become less intractable.

Nor is the argument about limitations on historical knowledge really convincing. No doubt much valuable information has been lost. It is clear enough, however, that historical materials that lend themselves to quantitative research, even if they do not cover everything, are enormously abundant. Some great storehouses of information such as census records and tax records are still relatively unused, except by a few pioneers. Other rich sources such as recorded votes in legislative bodies have been used only in a desultory and sporadic fashion, and much more could be done with them. Ample materials exist for collective biographies of groups of prominent individuals, and in some cases obscure ones too; for the economic and demographic characterization of constituencies; and for ascertaining the relationship of the facts unearthed in such investigations to political choice. Evidence is particularly rich for social and political history, two areas in which quantitative methods have not been extensively attempted.

Furthermore, it has proved possible, again and again, to describe in quantitative terms things that were formerly thought to fall beyond the reach of this net. Matters that seemed to an earlier generation unquantifiable can sometimes be caught and measured by a change in approach or by reaching a clearer perspective concerning what it may be most profitable to measure.

28 An extended account of the work that has been done and that might be attempted along this line in American history has been given by Samuel P. Hays in "New Possibilities for American Political History: The Social Analysis of Political Life," a paper presented at the Annual Meeting of the American Historical Association in December 1964; see also id., "Archival Sources for American Political History," American Archivist, XXVIII (Jan. 1965), 17–30.
This applies, for example, to the study of attitudes, a field in which notable advances have been made over the last several decades. David Hume, speaking through the mouth of Philo, a man of “careless scepticism,” argued that “controversies concerning the degrees of any quality or circumstance” can never “reach a reasonable certainty or precision.” Thus, he says, it is impossible to settle how great a general Hannibal was or “what epithet of praise Livy or Thucydides is entitled to . . . because the degrees of these qualities are not, like quantity or number, susceptible of any exact mensuration, which may be the standard in the controversy.”

Even if we cannot measure qualities of excellence, however, we can perfectly well measure opinions about them, which are all we have to go on anyway, and this is done all the time with questionnaires. Similarly, ways have been found to measure degrees of liberalism and conservatism by indexes in which men have come to place some reliance, or degrees of attachment to a particular cause or principle, or degrees of interest or apathy regarding political questions, or even degrees of patient welfare in a hospital. It has been possible to do this last by a set of objective tests that fit into a cumulative scale and that have turned out to be reliable and consistent.

Recent quantitative research in history contains several examples of a tour de force of this kind, attempts—fairly convincing attempts—to measure what previously seemed impossible to measure. One is the effort of Alfred H. Conrad and John R. Meyer to appraise the profitability of slavery and the efficiency of the slave labor market in the American South before the Civil War. It would be difficult to summarize here their complex and rather technical analysis, but it is interesting that their conclusions tell strongly against the long-standing though not wholly unchallenged view that the system of slavery was being undermined because of its unprofitability and because of the impossibility of maintaining and allocating a slave labor force. They found, on the contrary, that “slavery was apparently about as remunerative as alternative employments to which slave capital might have been put” and that: “Slavery was profitable to the whole South, the continuing demand for labor in the Cotton Belt ensuring returns to the breeding operation on the less productive land in the seaboard and border states.”

Another example is the attempt by McCormick to describe, for the period in which he does research, the relation between the economic status of members of the

---

24 David Hume, *Dialogues concerning Natural Religion* (London, 1779), Pt. XII.
electorate and their political choice. This topic, though important, is difficult for historians to study since the rich and poor in an electorate are generally buried in the anonymity of mass figures, and it is now virtually impossible to distinguish who voted for whom. McCormick, however, in attacking the problem, was able to take advantage of the dual franchise existing in North Carolina in the years 1836-1856, when only adult freemen who owned fifty acres of land within the county could vote for a member of the state senate while all freemen, including the above, who had paid county or state taxes, could vote for governor. By comparing, county by county, the size of the vote cast for governor with the comparable vote for state senators it was possible to determine the proportion of the electorate that could not meet the fifty-acre requirement. Then, by examining the distribution of each class of the electorate between the two major parties, McCormick was able to reach some conclusions on the relation of economic status to party affiliation. His finding, one of considerable interest, was that “the economic distinction implicit in the dual suffrage system had no substantial significance as a factor in determining party alignments in these North Carolina elections.”

In any case, the complexity and the limited scope of historical information are not arguments against quantification in particular. These limitations exist no matter what techniques are used. They arise from the insufficiency of the evidence and not from the peculiarities of the method. The obstacles to quantitative generalizations apply with equal force to nonquantitative ones, and what cannot be done with statistics cannot be done without them, either. No serious student of methodology would contend that a disciplined approach can overcome the inherent frailties in the data. But it hardly follows that, when the sources are suspect or the facts incomplete, an impressionistic, subjective approach can surmount these difficulties. Problems due to inadequacy of the data may be brought out more sharply and may become more apparent in a formal and systematic investigation, but they cannot in any circumstances be evaded.

(3) The objection is also sometimes made that the general conclusions of a quantitative investigation are not proved by the figures. This is, of course, true, and no one who knows anything about statistical theory would argue otherwise. To expect finality for the broader conclusions of a quantitative investigation is to misconstrue the nature of the approach. On this

point some misunderstanding apparently exists for, in everyday speech, reckless claims are sometimes made as to what "statistics prove." Actually the range of statistical proof is limited. A statistical table is nothing more than a convenient arrangement of the evidence, and it proves only what it contains: that there was, for example, a relationship or, more usual, a partial relationship between two variables. Theories that attempt to account for such a relationship, in the sense of fitting the findings into a wider conceptual scheme, are not proved by the figures. They are merely propositions that appear to explain what is known in a plausible fashion and that do not conflict with any relevant evidence that can, after a conscientious search, be uncovered. This is not to say that they are nonsense, for they may be supported by persuasive arguments. Yet since, notoriously, different arguments have proved persuasive to different audiences, the broader inferences from a quantitative investigation can scarcely be accepted as final. Thus it is possible, if the information is available, to establish how people voted, but it is much more difficult to say why they voted as they did. There might, of course, even be some difference of opinion on how they voted: for example, the accuracy of the records or of the tabulations made from them might be challenged. Yet such a disagreement is clearly on a lower level than a disagreement about men's motives, and there is a greater likelihood that it could be resolved through collecting and arranging the relevant data by acceptable procedures. In regard to more general explanatory propositions, however, a statistical inquiry, like any other method of verification, can only disprove. If the hypothesis does not fit the evidence, it may be rejected; in this sense a quantitative finding can indeed be conclusive, "Once we recognize that the Jacksonians won either by narrow majorities before 1837 or by narrow pluralities after that date, or frequently failed to win by any margin, it will surely become apparent that there is no basis for explanations that tell why they were the 'popular party.'"28 The absence of unfavorable findings does not, however, prove an explanatory generalization for there may be some other explanation, and it is also possible that adverse evidence may be discovered later. Strictly speaking, a generalization of this kind is never proved and remains on probation indefinitely.29

Hence there is always, in quantitative research of any scope, a gap between observation and theory. To bridge this gap it may be necessary to resort to assumptions that are not demonstrated by the evidence. Some recent

28 Benson, *Concept of Jacksonian Democracy,* 289–90.
presentations of this kind depend not only on the figures but also on the use of hypotheses that are designed to show either what the figures measure or what their relevance is to certain general questions. The gap cannot always be bridged. Interesting findings may be obtained that are difficult to explain in the sense of devising an acceptable theory that will account for them. An example from my own research is a cumulative scale, derived through Louis Guttman's technique of scalogram analysis, that ties together votes in the House of Commons on a number of different subjects in a way consistent with the hypothesis that they all measure a single variable. Yet the nature of this variable, this larger issue that subsumes many smaller ones, has proved difficult to determine. Though the existence of the scale can be demonstrated with about as much certainty as can ever be obtained in historical research, the characterization of it can be, for the present, only tentative and hypothetical. Comparable dilemmas have sometimes been encountered in other fields.

The hypotheses used to connect observation and theory are, no matter how plausible they at first appear, always open to challenge. The broader conclusions of a quantitative presentation may be vulnerable regardless of the accuracy of the mathematics or the reliability of the original data, and questions may be raised about them that cannot be answered by a resort to numbers. A critic may accept the findings, but then point out that the conclusions based upon them follow only if certain assumptions are made, and go on to question these assumptions. Some of the large modern quantitative studies have been criticized exactly on this ground: that the chain of argument, the series of connecting hypotheses, was too long and too tenuous to make the conclusions convincing.

By the same token, a quantitative investigation may not and often will not settle an argument. It may settle certain disputed points about the evidence. The discussion of larger questions of historical interpretation, however, concerns not merely what the facts were but also what may be inferred from them, and on this level controversy may continue. A quantitative finding may be open to more than one interpretation: in some cases it can

---

80 Aydelotte, "Voting Patterns," 148-51.
81 "In spite of the great social and scientific usefulness of psychological tests it must be acknowledged that for the most part we have had very inadequate ideas as to what it is that they actually measure." (Joy Paul Guilford, *Psychometric Methods* [2d ed., New York, 1954], 476.)
83 Greer based his conclusions on a calculation of what percentage each social or occupational group constituted of the total number of victims of the Terror. If he had argued from percentages figured in the other direction—designed to show what proportion of each of the various divisions of French society was executed in the Terror—the picture would have
be used to support either of two alternative and mutually exclusive theoretical schemes. It can also happen that quantitative results that appear to disprove an accepted theory will simply be “explained away.” This procedure can be quite legitimate, since it may prove possible to achieve a reformulation of the earlier view, which preserves some of the original insights, but does not conflict with the new evidence. If contradictory findings continue to accumulate, however, it may eventually be more satisfactory to abandon the earlier position altogether.

Quantitative procedures by no means preclude, nor indeed can they possibly eliminate, the use of value judgments, speculations, intelligent guesses, or “the imagination and intuitive feel which the historian, and for that matter the social scientist, should bring to his subject.” What is gained by attempting such exactitude as the circumstances allow is not finality but reasonable credibility, not the elimination of subjective factors but the minimizing of their role. No greater claim than this would be asserted by responsible social scientists or statisticians.

These points, though they are elementary, are not always understood or remembered. Quantitative findings are impressive in appearance and may, by their psychological impact, numb or blunt the critical abilities of the reader. It does occasionally seem to happen that a statistical presentation wins acceptance not through intellectual persuasion but through a kind of hypnosis. There is, however, no magic about quantitative evidence. It may be more conveniently arranged and, on the points it covers, more complete than other forms of documentation. Its significance, however, depends on what can be inferred from it, and such inferences, like all other inferences, may be fallible.

In case this appears puzzling, it may be helpful to summarize the hypothetical illustration given by Hans Zeisel. If Company A increases its sales volume in a year from one to two million dollars and Company B, a bigger outfit to begin with, increases its sales in the same period from four to seven million dollars, then one could argue either that Company B did better since its net increase was three times that of A, or that Company A did better since it increased its sales 100 per cent in comparison to B’s 75 per cent. Which alternative is preferred depends not on the figures but on what causal assumptions are implied in making the comparison and on what kinds of questions the investigator wishes to test. (Hans Zeisel, Say It with Figures [4th ed., New York, 1957], 8-13.)

This disadvantage is not, of course, peculiar to quantitative procedures. On the contrary, quantitative evidence stands, in this respect, on a level with all other kinds of evidence, and arguing from it is subject to the same rules and the same hazards. The danger of false reasoning from good evidence occurs in any kind of research. It is not only in the field of statistics that men may agree on the facts but disagree on the inferences to be drawn from them.

(4) Questions have been raised not merely regarding the feasibility and reliability of quantitative research in history but also regarding its usefulness. It is sometimes argued that quantitative findings, even if they can be trusted, tend to be trivial, inconsequential, and uninteresting. This is because any system of classification, such as is needed for such work, uses only a small part of the available information and leaves out the full richness of reality. Hence the ordinary statistical categories are too crude and threadbare to explain the complicated chains of events with which history is concerned. The problems in which historians are most interested are so complex that they elude these methods. One critic holds that: "almost all important questions are important precisely because they are not susceptible to quantitative answers."86

It is true, of course, that any quantitative procedure involves using only selected classes of data. It is seldom possible to include everything, or to come anywhere near this. Hence, statistical tables, though they seem impressive, may also present an appearance of bleakness or barrenness which can act as an impediment to thought. Often they will not stimulate the imagination as the detailed recital of an individual case will do. Indeed, it is useful, when one comes to an impasse in interpreting the figures, to turn to the consideration of individuals about whom much is known. Such individuals may not be representative, and one cannot generalize from them to the whole group; a study of them may, however, yield suggestions or leads, fresh hypotheses that can be tested, which will make the evidence as a whole more intelligible. It is always necessary, when working with the figures, to remember that they do not tell the whole story, that many elements of the situation are not reflected in them, and that what they do not cover may turn out to be more important than what they include. To interpret the quantitative evidence it is generally necessary to have recourse to the more conventional sources of historical information: memoirs and biographies, congressional debates, private papers, and the like.

The charge that quantification abstracts and uses only limited parts of the available information, however, is not an objection to this method specifically. On the contrary, any generalization abstracts. A generalization is a comparison of a number of cases, not in terms of all the attributes of each, but in terms of certain selected attributes in respect to which the cases are comparable. This problem is not peculiar to quantification; it arises in any research in which a conscientious effort is made to substantiate general statements.

The objection that the findings of quantitative studies are not significant sometimes takes other forms. It has been alleged, for example, that this kind of research is destructive and not constructive and that: "the recent use of quantitative methods to test historical generalization has resulted in the wholesale destruction of categories that previously held sway in the historian's vocabulary without supplanting them with new generalizations of comparable significance."\(^{39}\) As an objection to quantification, however, this argument has no weight for it applies equally to any form of verification. All verification is in this sense negative. The argument fails to distinguish between the two quite different activities involved in research: getting ideas and testing them. Quantitative inquiries are generally directed to testing hypotheses formulated in advance. It has frequently been observed that, in work of this kind, a flat-footed empiricism is not likely to rise above a fairly low conceptual level and that systematic thought will progress more rapidly when it is directed by some adequate general hypothesis. The point should not be pushed too far for it occasionally happens that important relationships are not anticipated, but emerge as windfalls after the inquiry is completed. Also, in an area in which little work has been done, the original investigations must often be to some extent exploratory. It would be pedantic to insist on a full-fledged hypothesis in every case.\(^{38}\) Nevertheless, the criticism that quantitative methods destroy and do not create is clearly based on a mistaken notion of the usual role of hypothesis in research. Hypotheses and generalizations are not simple inductions that emerge of their own accord from the evidence; they have, as is now better understood, different and more complex origins.\(^{39}\)


\(^{39}\) I have discussed this point at greater length in "Notes on the Problem of Historical Generalization," in Generalization in the Writing of History, ed. Gottschalk, 163-72.
Nor does a negative finding necessarily represent a dead end. If a generalization is wrong, it is useful to have it disproved; the disproof constitutes an advance in knowledge. As J. H. Hexter observes: "it may be worth saying that violent destruction is not necessarily of itself worthless and futile. Even though it leaves doubts about the right road for London, it helps if someone rips up, however violently, a 'To London' sign on the Dover cliffs pointing south."\(^{40}\) A negative finding can be, in some cases, as valuable as a positive finding, depending on what theoretical inferences follow from it. Furthermore, to blame the quantitative method for disproving bad hypotheses is to blame the doctor instead of the disease. What is at fault is the mistaken opinion, not the technique that reveals when we have gone astray. The remedy is not to abandon the technique but to try to develop a new theory that fits the evidence better.

It is also sometimes argued that quantitative methods only prove the obvious, that they merely demonstrate, by an unnecessarily cumbersome apparatus, what everyone already knew.\(^{41}\) It is admitted that they can occasionally be used to disprove certain crude generalizations that still appear in the textbooks. Yet, it is said, the crudity of such generalizations is already widely appreciated, and, on the whole, they are not accepted by sophisticated historians. In other words, quantitative techniques are useful only when historians have made fools of themselves. Their function is to clear away rubbish. However, if there is no rubbish, if scholarship in a field has been reasonably careful and responsible, a quantitative analysis is unlikely to reveal anything that is not already fairly well understood.

This criticism, also, is not well taken. Even if research merely confirms in a more conclusive fashion what some people already believe, it is good to have this additional assurance and to establish this belief on a more solid foundation. Also, on many questions that can be studied by quantitative methods, the answer is by no means a matter of course. More often there is evidence pointing in both directions, and both sides of the argument have been supported with some plausibility by different individuals. In such cases, it is useful to establish which of two contradictory statements comes closer to describing the total evidence and just how close it comes. It might be added that, in disputes of this kind, either answer will be "obvious" in the sense of being already familiar, even though the two alternative answers exclude each other. Furthermore, the results of quantitative investigations have fre-

Several examples have already been given; another is Fogel’s attempt to appraise the role of the railroads in American economic growth, which resulted in the conclusions, disheartening to some enthusiasts, that even in the absence of railroads the prairies would have been settled and exploited, that the combination of wagon and water transportation could have provided a relatively good substitute for the railroad, and that “no single innovation was vital for economic growth during the nineteenth century.”

Whether the results of a quantitative investigation are important or trivial is and can only be a matter of opinion. The presumption of significance is based not on a demonstration of fact but on a judgment of value. This applies, incidentally, even to the so-called “tests of significance” commonly used in statistics. Properly speaking, they are evaluations of probability, and, while probability can be mathematically determined, the degree of it that will be regarded as acceptable in any study is a question not of mathematics but of the investigator’s preference. A quantitative study, it might be said, is significant if the investigator thinks it is and can persuade others to share his view. Speaking simply on this basis, it seems difficult to support the assertion that the topics open to quantitative investigation are of no consequence. Far from this being the case, an intelligent use of the method opens up a host of new, potentially interesting questions that could be approached in no other way. Some of these possibilities and some of the studies conducted along these lines have already been discussed. Perhaps it is enough to say here that the substantive weight of the findings of the limited number of historians who have attempted quantitative research is already impressive enough to render the accusation of triviality something less than plausible.

It seems reasonable to argue, furthermore, that the significance of a project of research does not depend on whether it is quantitative or not. Quantitative presentations vary greatly in value. They may be significant or trivial, interesting or uninteresting, and it is incorrect to suppose that they are all on the same level in these respects. What gives them such worth as they may have are the importance of the problem, the abundance, reliability, and relevance of the available evidence, and, above all, the intelligence with which the work is executed.

In fact, what is most needed in research of this kind is not the automatic application of certain techniques but, rather, qualities of logic and imagination. The main problems here, as in all research, are not technical and mechanical but intellectual and analytical. It is not easy to make the figures

“talk” or to show their bearing on significant problems, and nothing is
drearier than a presentation that merely summarizes the evidence. I
am disturbed by students who want to do quantitative research and who
seem to expect that this will solve their problems and that the application
of a method will save them the trouble of thinking. This expectation is
erroneous. Quantitative techniques, though they may play a crucial role in
demolishing previous theories, are usually not adequate, by themselves, to
establish general alternative hypotheses. They are nothing more than a means
of deploying the evidence, although they perform this limited service won­
derfully well. Once this subordinate and ancillary work has been done, how­
ever, the basic problems of historical interpretation still remain to be dealt
with; they are not to be resolved by a gimmick. The greatest hazard in
quantitative research is not that of neglecting techniques but that of becom­
ing too much absorbed in them. This danger is particularly threatening now
because of the rapid development of mechanical facilities for the processing
of data. It is only too easy to become absorbed in the gadgets and to forget
the ideas. The refinement and sophistication of methods, though desirable
in themselves, can become a kind of escapism, an evasion or postponement
of the intellectual tasks that must ultimately be faced.

In general, the discussion of quantification in history has involved much
talking at cross-purposes. Many of the common objections to this approach
seem to arise from a misconception of its function. They appear to assume
that claims have been made for it that no responsible statistician would
make. No one well versed in this line of work would argue that all his­
torical materials can be quantified, that the figures provide any final demon­
stration of the broader inferences derived from them, or that the figures tell
the whole story. Such assertions are clearly improper. If they are not made,
however, as by informed workers in this line they are not, much of the
current offensive against quantitative techniques fails. The central point
around which discussion of the subject has in part revolved is not an intel­
lectual issue but a problem of communication.

The use of quantitative methods for history presents substantial difficul­
ties not always appreciated by enthusiasts or neophytes. Those who have
employed them are likely to be less starry-eyed about their possibilities than
those who have merely commended them without trying them. Indeed,
quantitative projects may be more glamorous in the planning stage than
they are after the results have been gathered; the findings sometimes turn
out to be flatter and less revolutionary than had been hoped.

Though the difficulties are real enough, however, it is not clear that they
constitute objections specifically to a quantitative approach, or that they can be resolved by dispensing with it. The standard objections are misconceived or placed out of context when presented as grounds for rejecting these methods altogether. Properly understood, these reservations serve not to discredit quantification but to mark the boundaries of what it can accomplish. Indeed, the apparent disadvantages of quantitative research, the impediments to generalization that it presents, are actually advantages for they call attention to limits in knowledge or to flaws in reasoning that might not otherwise be perceived or fully appreciated. When all reservations have been made, quantification has still shown itself, in the light of the considerable experience we now have, to be a powerful tool in historical analysis. It helps to make the work both easier and more reliable, and, in some cases, it provides a means of dealing with questions that could not be attacked in any other way. Those wrestling with problems for which this approach is appropriate can ill afford to dispense with it. In the general intellectual twilight in which historians are condemned to spend their lives, even some small effort to render the darkness less opaque may be advantageous.