



Aspects of Quantitative Research in Economic History

Lance E. Davis; Jonathan R. T. Hughes; Stanley Reiter

The Journal of Economic History, Vol. 20, No. 4 (Dec., 1960), 539-547.

Stable URL:

<http://links.jstor.org/sici?sici=0022-0507%28196012%2920%3A4%3C539%3AAOQRIE%3E2.0.CO%3B2-J>

The Journal of Economic History is currently published by Economic History Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/eha.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Aspects of Quantitative Research in Economic History

I

IF we are successfully to relate our work with the main body of Economic History, we must be able to show the fundamental relationship between quantitative analysis and more conventional methods of economic historians. The historian reconstructs events of the past, and with them attempts to understand the institutions and modes of behavior associated with those events. He seeks to construct a consistent story revealing the fundamental nature and meaning to us of the past, thus creating insight into the past and understanding of it—something considerably beyond a mere account of what probably happened. However, this story must be based upon, and be consistent with, the reconstructed events of the past, “what probably happened.”

This view of historical study is a familiar one and, in fact, almost any working historian would accept it as a definition of his activities. What has this view to do with quantitative methods in economic history, and how does quantitative economic history differ from non-quantitative writing in economic history? The answer lies partly in the nature of the materials from which the reconstruction is to be made, and partly in the technique employed to analyze these materials. We may distinguish two broad classes of materials. First, each society generates its own accounts of, and commentaries upon, contemporary affairs, as well as its own histories of the past. This class of materials can be more or less explicitly labeled historical writings. Second, in the course of its characteristic processes each society generates a body of artifacts, debris left behind in time. The processes of economic life, for example, produce masses of receipts, books of account, legal documents, tax returns, and various kinds of rolls, lists, and records; and these materials can be used to reconstruct the past. Such a reconstruction is analogous to the process frequently employed by archeologists who utilize the surviving debris of an ancient city to provide materials for understanding the civilization that built it and lived there. Characteristically, this material consists of a mass of individual items, each of which contains a relatively insignificant piece of information. This collection, originally generated for purposes other than historical study, usually requires

reorganization and analysis to enable the information contained to be brought to a form useful to the historian. That is, the information must be organized on some principle and made the basis of inferences about the past, a task familiar to the historian, but that also contains the essential elements of a statistical problem.

If we had at hand every item of the kind under consideration (for example, every warehouse receipt issued in New York State between 1870-1900), we would face the task of formulating meaningful questions with which to confront these data, and of organizing the data to bear on these questions. Thus we might be interested in the geographical distribution of warehouses, in the time-shape of business activities, in the commodity composition of consumption, or in the profitability of warehousing, and each question would require a different method of systemization and analysis. We must, therefore, first formulate a statistical-historical model in terms of which historical meaning can be given to the observations. Having formulated such a model, we are still confronted with a large collection of observations that must be statistically summarized and described. We shall have more to say later about the problems of data processing.

In fact, however, since we generally do not have the complete collection of all observations of a given land, but rather only a sample of the surviving ones, the study of the historical process is even more complicated. For in addition to the question of analysis and systemization, we must also decide, on the basis of the observations at hand, what the whole collection would reveal if we had access to complete information. This last is the problem of statistical inference.

In brief, the logical structure necessary to make historical reconstructions from the surviving debris of past economic life essentially involves ideas of history, economics and statistics. The offspring of such an act of interdisciplinary miscegenation calls for a name worthy of it; at Purdue the resulting discipline has been labeled "Cliometrics."

II

Your chairman has on occasion jocularly referred to us as the authors of *La Loi Lafayette* because of the frequent appearance of the tools of statistics and data processing in our work; however, this classification gives us too much credit. An examination of the literature of economic history indicates that, while the qualitative stream in the discipline has usually been the larger, there has been from earliest times a significant

and respectable flow of quantitative work. The political arithmeticians—Graunt, King and the like—as early as the seventeenth century were trying to infer from data an explanation of some aspects of economic history.¹ In 1851 William Newmarch produced his pathbreaking study on the circulation of bills of exchange—a study that, in statistical sophistication (given the knowledge of the time), is the equal of anything produced more recently.² At yet a later date, F. W. Taussig's *History of the Tariff*, and still later, the work of Arthur Cole (as represented both by his work on the evolution of the American foreign exchange market and with W. B. Smith, *Fluctuations in American Business 1790-1860*), were attempts to infer the state of the world of the past from quantitative information.³ The National Bureau of Economic Research, although not much given to interpretation, has produced a vast amount of quantitative information that others could use, and the works of some of their authors (Kuznets, for example) certainly fall within the category of quantitative economic history.⁴ More recently excellent work combining economic theory and quantitative methods can be found in R. C. O. Matthews's *A Study in Trade Cycle History* and in what is perhaps the most notable recent study of this kind, the paper by Conrad and Meyer on slavery in the ante bellum south.⁵

¹ See, for example, Charles D'Avenant, *An Essay Upon the Probable Method of Making a People Gainers in the Balance of Trade* (London, 1699); John Graunt, *Natural and Political Observations Mentioned in a Following Index and Made Upon the Bills of Mortality* (London, 1662); Gregory King, *Natural and Political Observations and Conclusions Upon the State and Condition of England 1696* (London, 1810); Charles H. Hull, ed., *The Economic Writings of Sir William Petty* (Cambridge, Mass., 1899).

² William Newmarch, "An Attempt to Ascertain the Magnitude and Fluctuations of the Amount of Bills of Exchange (Inland and Foreign) in Circulation at One Time in Great Britain, in England, in Scotland, in Lancashire, and in Cheshire, Respectively, During Each of the Twenty Years 1828-1947, Both Inclusive; and Also Embracing in the Inquiry Bills Drawn Upon Foreign Countries," *Journal of the Statistical Society of London*, XIV (1851), 143-92.

³ Frank W. Taussig, *The Tariff History of the United States* (New York and London; G. P. Putnam's Sons, 1888); Walter B. Smith and Arthur H. Cole, *Fluctuations in American Business 1790-1860* (Cambridge, Mass.: Harvard University Press, 1935); Arthur H. Cole, "Seasonal Variation in Sterling Exchange," *Journal of Economic and Business History*, II (Nov. 1929) 203-318; Arthur H. Cole, "Evolution of the Foreign Exchange Market of the United States," *Journal of Economic and Business History*, I (1928), 384-421.

⁴ See, for example, Simon Kuznets, *National Product Since 1869* (New York: National Bureau of Economic Research, 1946); or Simon Kuznets, *Secular Movements in Production and Prices; Their Nature and Bearing Upon Cyclical Fluctuations* (Boston and New York: Houghton Mifflin, 1930); or "Long-Term Changes in the National Income of the United States Since 1870," *Income and Wealth of the United States Trends and Structure*, Series II, pp. 10-246.

⁵ Robert C. O. Matthews, *A Study in Trade Cycle History; Economic Fluctuations in Great Britain, 1833-1842* (Cambridge, England: University Press, 1954); Alfred H. Conrad and John R. Meyer, "The Economics of Slavery in the Ante Bellum South," *Journal of Political Economy*, LXVI (April 1958), 95-130.

None the less, the total amount of work in the field is small. Why? Is it because quantitative work is unrewarding? We think not. The dearth of quantitative economic history can probably be traced to two factors: first, the extraordinary effort that has been necessary in the past to sift and classify quantitative information; and second, the relatively recent development of statistical theory and techniques capable of handling these problems. Let us examine each of these problems in turn.

III

Recently developed computing equipment has opened to economic historians the possibility of performing prodigies of data-processing and statistical calculation. Where the archeologist can dig with spade and hard labor into the mounds of the past to unearth artifacts, the economic historian needs power shovels and bulldozers to move the mountains of paper records. The power shovels and bulldozers are now available to economic historians, but these developments in the analysis of data by machine processing methods have scarcely been applied to the more significant questions of economic history. We need to recognize that data-processing at last provides us an opportunity to study the kinds of problems that, because of the unwieldiness of the purely mechanical processes of computation, have long been cast aside.

There is nothing novel or revolutionary in the problems themselves. There have always been problems in which masses of data had to be "processed" by one means or another; the early work on index numbers by Jevons and Sauerbeck are well-known examples of such large-scale computations. Another, and better, example of early data-processing is to be found in the work of one of our most illustrious predecessors, William Newmarch. His work is a particularly germane example of the nature of data-processing problems.

Newmarch's celebrated survey of the circulation of bills of exchange in Great Britain, one of the most brilliant contributions to the British monetary-policy debate of the mid-19th century (although not one which noticeably affected policy, interestingly enough), and subsequently, one of the pieces of historical evidence most often used in the study of 19th century monetary phenomena, appeared in 1851.⁶ Anyone who has processed data will experience a distinct "shock of recognition" when he first reads Newmarch's paper. Newmarch obtained a

⁶ Newmarch, "Circulation."

sample "at hazard" of the total bill circulation. The sample size was 4,367 inland and foreign bills with a nominal value of £1,216,974. From each bill he took three pieces of information, or a total of 13,101 separate pieces of numerical information. On the basis of these data and the stamp-tax returns on bills of exchange for Great Britain and certain subdivisions Newmarch was able to compute an estimate of the total bill circulation. In his computation only three arithmetic operations were required once the initial classification and organization of the data had been done. Such a study by present-day computing standards would not be cumbersome. In "The First 1,945 British Steamships," for example, there were 13 initial observations per ship, or 25,285 separate pieces of numerical information on 1,945 punch cards.⁷ Once on the cards, the actual computing (far more extensive than was Newmarch's) of tonnages, estimated speeds and the index of transport capacity was not a particularly tiresome job. But in Newmarch's case, one hundred years ago, 13,101 separate pieces of information to be classified, summed, averaged, and the results used to convert stamp-tax yields into aggregate value figures was an enormous undertaking. After explaining his ingenious methods to his audience, Newmarch noted that the final operations were:

. . . clearly a mere matter of calculation; but I confess that, if I had foreseen, before I undertook the task, the extent and severity of the labour it would impose, I am not at all certain that I should have ventured upon the inquiry.⁸

That data-processing on this scale was done so long ago with quill pens instead of electronic computers underscores our main point. It is only the absence of machinery which is novel in the Newmarch story, not the presence of a data-processing problem. Some kinds of historical problems are by nature data-processing problems. In the past they have too often (but perhaps understandably) been neglected simply because too much labor was involved. Today, with the purely mechanical computational problems much reduced, economic historians have the means to study these questions. In particular these new techniques will permit the opening of new sources that, while always in existence, have heretofore been largely closed to research because of the magnitude of the task involved. First, business records that hold many of the answers

⁷ Jonathan R. T. Hughes and Stanley Reiter, "The First 1,945 British Steamships," *Journal of the American Statistical Association*, LIII (June 1958), 360-81.

⁸ Newmarch, "Circulation," p. 149.

to questions concerning early 19th century American development can be made available to the study of broader questions than the history of a single business. While corporate letters, minutes and like documents have long been utilized in business history, the labor involved in organizing sales slips, time cards and the like has frequently prevented their use in producing data on prices, output, investment and employment that would be more useful to the economic historian. Second, largely unorganized government data existing in committee and bureau reports could be brought together in manageable shape. And third, even correspondence, long a bug-a-boo of historians, could, perhaps, be better analyzed by data-processing techniques together with some form of content analysis. Both economics and economic history stand to gain enormously if this work is done and done wisely.

Our work at Purdue along these lines is, we hope, only the beginning of extensive data-processing work in Economic History. Since 1957 five data-processing studies in Economic History have been produced at Purdue. In all of our studies we have developed entirely new statistical series which are now readily available as "building blocks" for other economic historians to use in their studies. Our main results from these five papers may be briefly summarized.

In "Sources of Industrial Finance: The American Textile Industry, A Case Study," one hundred seventy-five observations on each of eight financial variables were brought together in a multiple regression model. The analysis of this data indicated that heretofore economic historians had tended to overstate the role of retained earnings in 19th century American corporate finance and understate the role of borrowed capital. In addition, the analysis indicated, as might be suspected, that the importance of loans grew concomitantly with the development of the capital markets and that a firm's capital structure was responsive to short-term changes in the capital markets, output and employment.⁹

In "Stock Ownership in the Early New England Textile Industry," data-processing techniques were used to systematize some 3,782 separate stock accounts (representing the equity holdings of a least 854 individuals and firms in eleven textile mills) to uncover the trends in stock ownership over the period 1829 to 1859. The analysis indicates, first, that while mercantile capital represented a large proportion of the investment, the transfer from mercantile to industrial capital appears to have been slower than has been generally assumed. Second, that there

⁹ Lance E. Davis, "Sources of Industrial Finance: The American Textile Industry, A Case Study," *Explorations in Entrepreneurial History*, IX (April 1957), 189-203.

was a considerable degree of backward integration with textile merchants and mercantile firms contributing a significant block of capital. Third, that financial intermediaries made a substantial contribution to the finance of new industry. And fourth, that out-of-state and foreign investors made no significant investments.¹⁰

In "The First 1,945 British Steamships," the growth of the British steam merchant marine up to 1860 was chronicled and measured in detail, by type of propulsion and build. It was shown that not only was this development more rapid than had previously been thought, but that the transport capacity of the fleet grew more rapidly than did the tonnages to the extent that it could have played the powerful role in overseas earnings which had been assigned to it by contemporaries, and which had been heavily discounted by modern scholars.¹¹

In "The New England Textile Mills and the Capital Markets: A Study of Industrial Borrowing 1840-1860," 2,385 industrial loans were systematized by machine techniques in order to provide some new information about the ante-bellum capital markets. The analysis produced a new series of interest rates independent of the frequently-cited Bigelow estimates. It seemed to indicate that a theory of a sectored money market better explains the term structure of rates than does the more classical Lutz-Hicks expectation theory, and it presents some new data of the relative importance of various types of lenders in the markets and the cyclic and secular changes in the composition of this lender group.¹²

Finally, in "A Dollar Sterling Exchange 1803-1895," data processing techniques permitted us to organize 2,789 bills of exchange and, from the bill prices, to deduce a series of pure exchange rates. In addition to this new series, the analysis indicates that exchange rate stability did not always characterize "the gold standard"; instead, it is only after 1875, when transportation, communications, and the money markets had evolved into near-modern form, that exchange stability became common. Finally, the paper casts further light on the development of the foreign exchange market and the rise of the major foreign exchange houses—a subject previously explored by Arthur Cole.¹³

¹⁰ Lance E. Davis, "Stock Ownership in the Early New England Textile Industry," *The Business History Review*, XXXII (Summer 1958), 204-22.

¹¹ Hughes and Reiter, "1,945 Steamships."

¹² Lance E. Davis, "The New England Textile Mills and the Capital Markets: A Study of Industrial Borrowing, 1840-1860," *The Journal of Economic History* XX (March 1960), 1-30.

¹³ Lance E. Davis and Jonathan R. T. Hughes, "A Dollar Sterling Exchange 1803-1895," *Economic History Review* (August 1960).

IV

We now turn our attention briefly to statistical inference. Broadly speaking, statistical inference refers to a body of techniques that permit the user to garner with some confidence a knowledge of certain characteristics of populations on the basis of observations from these populations. For example, Newmarch's estimate of the average value of all bills based on the average of the sample drawn is a statistical inference. In that case the unknown quantity (called a parameter) is the average of all bills. In the Hughes and Reiter paper on the first 1,945 steamships a more elaborate, but essentially similar, technique was employed to estimate the speed and carrying capacity of the ships from a knowledge of other related facts about them.

In other cases the unknown quantity of interest might be, for example, a measure of dispersion, or the largest observation in the population. Inference of this kind can achieve remarkable accuracy, more accuracy sometimes than that yielded by an attempt to count an entire population. This was the case in World War II when Allied estimates, based on the serial numbers of samples of German military equipment observed in battle, proved to be more accurate than the information supplied to the German government from production records, and, in addition, were available much sooner than the data derived from the latter source.¹⁴ Further, there are techniques for the study of relationships among various observable quantities, ranging from simple regression analysis to highly elaborate statistical schemes for detecting the presence of association between one pair of variables in the presence of many other associations or influences. A simple multivariable regression model, for example, was used in the study of textile financing previously cited to sort out the effects on corporate finance of firm age and historic time.

Before closing this desperately brief discussion of statistical inference, we should point out that one ought not to have too narrow a concept of the phenomena subject to quantitative analysis. It is, of course, obvious that observations, given in the form of numerical quantities such as money amounts or physical units of output, are subject to quantitative analysis. However, the possibilities of quantification go much further than this. Any phenomena whose occurrence may be noted or

¹⁴ W. Allen Wallis and Harry V. Roberts, *Statistics, A New Approach* (Glencoe, Illinois: The Free Press, 1956), p. 20.

counted is quantifiable, and, in addition, coding sometimes permits apparently non-quantitative phenomena to be quantified. Thus attitudes, opinions and perceptions are studied with the aid of quantitative techniques by psychologists. Content analysis of written material provides another example of the usefulness of an extended notion of quantification. In our study of 19th century exchange rates we coded proper names, origins and points of payment of the bills. Thus "quantified," the mass of information could be handled and yielded us important evidence of the development of specialization in the exchange market.

V

We are not suggesting in this paper that there is to be a "new" economic history which will render non-quantitative economic historians technologically unemployed. It should be obvious that we regard ideas from statistics and data-processing as natural aspects of problems of historical study. It should also be obvious that the historian's special knowledge and viewpoint is essential to the useful employment of quantitative methods. Our main point is that modern statistical techniques and computing equipment make possible the intensive exploitation of a vein of historical materials that was perforce only little worked in the past; and that if even a few economic historians would take the time to learn even a little of these new techniques, the 1960's could easily prove the most productive years in the history of the discipline. On the other hand, if the discipline chooses to remain completely in the literary tradition, we can see small hope for anything but a continual rehashing of the already existing sources and a continuation of the century-long cleavage between economics and economic history—a cleavage that should soon disappear if the economic historian is able to provide the economists with new data and new interpretations of the process of economic life.

LANCE E. DAVIS, JONATHAN R. T. HUGHES AND STANLEY REITER,
Purdue University