

## The Computer and the Colonial Treasury Accounts: A Proposal for a Methodology

PAUL E. HOFFMAN\*

A WELL-KNOWN WEAKNESS of history-writing about the colonial period of Latin American history is the dearth of statistical information on all aspects of life in that period. Paradoxically there exist great quantities of data in the Archivo General de Indias and various national archives in Latin America from which statistical series could be constructed. In the AGI, for example, quantifiable data are to be found in most of the 1,953 legajos of the Contaduría section, and in perhaps a fourth of the 5,873 legajos of the Contratación section. Scattered throughout other sections are numerous reports on population, volume of trade, prices, and other statistically treatable aspects of colonial life.

Particularly important sources of quantifiable data are the nearly complete accounts of the various Spanish royal treasuries. Expenditures recorded in these accounts can shed light on royal policy and on many aspects of colonial life.<sup>1</sup> For example, the author's preliminary work with defense spending has shown that average annual expenditures for the decade 1565-1574 are at least eighteen times as much as for the decade 1535-1544. Inflationary factors alone would only double the level of expenditure. The remaining increase (1700%) represents changes in the corsair problem and a shift in the Crown's role in defense.<sup>2</sup>

\* The author is Assistant Professor of History at the University of Wyoming.

<sup>1</sup> The value of the royal treasury records for socio-economic history was first suggested by Clarence H. Haring, "Ledgers of the Royal Treasuries in Spanish America in the Sixteenth Century," *HAHR*, II (May 1919), 173-187. María Encarnación Rodríguez Vicente, "La Contabilidad virreinal como fuente histórica," *Anuario de Estudios Americanos*, XXV (1967), 1523-1542, surveys some of the recent literature based on these accounts and the problems of using them. Ismael Sánchez-Bella, *La organización financiera de las Indias, Siglo XVI* (Seville, 1968), also notes the need for studies of expenditures (cf. note 194, p. 250), as does José María de la Peña in *Archivo General de Indias de Sevilla, Gota del Visitante* (Madrid, 1958), 88-89.

<sup>2</sup> The figures are: 1535-1544, an average of 7,000 ducats per year; 1565-1574, an average of 131,000 ducats per year. These figures are based on materials

Heretofore the chief difficulty preventing scholars from using this mass of data on Spanish royal spending has been the lack of a system which would allow the rapid transformation of raw data into yearly or decennial totals and time series. Even the shortest run of statistics compiled by the use of note cards and an adding machine requires more time than the historian usually feels he can spend on that aspect of his study. Statistical studies covering centuries have involved expenditures of time and energy which few researchers can sustain.<sup>3</sup>

But the scholar need no longer rely on the slow method of data processing by note card and adding machine. Since the advent of high-speed digital computers there is no methodological reason to shrink from far-ranging statistical studies. Nonetheless, many historians are reluctant to use "the machine" because they fear that data so processed cannot be manipulated to remove various types of error or because they lack coding systems which will allow them to make their data machine-readable. This paper will suggest why this belief is largely mistaken. It will then examine coding systems which should enable scholars to study comprehensively the Crown's expenditures of its colonial revenues, as well as other aspects of colonial life which may be reflected in the treasury accounts.

The records of the royal treasury are notorious for the atmosphere of fraud which is said to surround them. An examination of the literature reveals that such statements usually refer to the income side of the ledger, where various forms of tax evasion make inferences hazardous about the economic state of the colonies.<sup>4</sup> It would be a mistake to assert that expenditures were not also falsified or

---

from the accounts of the Casa de Contratación, Santo Domingo, Havana, Tierra Firme, and Cartagena. They thus represent most but not all of the royal funds spent for defense. Paul E. Hoffman, "The Defense of the Indies, 1535-1574. A study in the Modernization of the Spanish State" (Ph.D. dissertation, University of Florida, 1969), 296-317. A revised and amplified study of defense costs is being prepared, using the coding system described in this paper.

<sup>3</sup> Earl J. Hamilton records that his study of *American Treasure and the Price Revolution in Spain, 1501-1650* (Cambridge, 1934; reprinted, New York, 1965) required over 43,250 hours of labor. Pierre and Huguette Chaunu spent over five years on *Séville et l'Atlantique* (11 vols., Paris, 1955-1959).

<sup>4</sup> For examples see the remarks in Chaunu, *Séville*, I, 97-110; and in Sánchez-Bella, *Organización*, 204-250. Besides tax evasion, the revenue side of the ledger involved frauds having to do with the collection of one type of money while crediting the accounts with another of lesser value, the collection of old debts in money worth less than that involved in the original debt, or the use of royal funds for private speculations before they were credited to the royal accounts. It is thus nearly impossible to know what should have been collected; the same is true of what was collected.

subject to frauds of various types. The number of such frauds was kept to a minimum, however, by several devices: the elaborate safeguards built into the cycle of draft (*libranza*), notarized payment (*carta de pago*), entry into the account books in the presence of several witnesses, and audit at one or more levels of the administration.<sup>5</sup> Those frauds which escaped detection in this process simply became costs of government so far as contemporaries were concerned. Thus, unlike frauds involving revenues, those undetected in the expenditure data do not prevent the historian from constructing statistical series whose figures accurately reflect the transactions which occurred.<sup>6</sup>

Besides the problem of fraud, there are other reasons for suspecting the reliability of statistics based on the accounts. It is well known that the estimates (*tanteos*) and summaries of the accounts (the latter appearing at the end) were often inaccurate. Sometimes the compilers made errors in arithmetic, or the royal officials tried to juggle the books and show them in a better light than might otherwise be the case. Moreover, the categories in which revenues and expenditures were classified are often of little interest to the modern historian, nor are categories always consistent. Not least of all, the accounting period was seldom the same as the calendar year. It might range from a few months to several years in length. Statistical series based on the summaries of the accounts are thus without a uniform temporal base. These problems of arithmetic accuracy, juggled books, poor classification, and a lack of uniformity in the accounting period became more common as the colonial era progressed. Small wonder that many historians view any statistics from this period with suspicion.

In order to obtain reliable statistics from accounts subject to these uncertainties, the researcher must approach them entry by entry, throwing out those entries which were "tested" or otherwise

<sup>5</sup> For details of this cycle see: Sánchez-Bella, *Organización*, 250-256. Sánchez-Bella cites the pertinent sources. An additional safeguard was contained in the accounts of the persons receiving the goods for use. Particularly under Philip II, these persons were held to strict accountability for all supplies received. The Royal Treasury was thus able to recoup some of its losses from misappropriation of supplies.

<sup>6</sup> It should be pointed out that absolute figures are not required for most problems. Figures of the correct order of magnitude which are accurate to three or four digits are generally as much as can be expected in any statistical study, including those done from twentieth-century sources. Figures based on the expenditure data from the royal accounts have this type of accuracy. The problem with the income side of the ledger is that it is almost impossible to determine whether the computed figures are accurate to four places and of the correct order of magnitude. The reasons for this appear in note 4.

questioned by the auditors, which are obvious duplications, or which are mentioned as fraudulent in other sources. Once disassembled and checked in this fashion, the accounts can be reassembled on the basis of the calendar year, according to categories of historical interest, and with arithmetic accuracy. The keys to these rectifications are individual data cards for each entry or item in it, a coding scheme or schemes, and some fairly simple computer programs. The only way in which this system differs from the manual method is that the chore of assembling the data is handled by machine, usually with considerable savings in time.<sup>7</sup> The historian, therefore, may have as much confidence in the results as if he had used some other method to manipulate the data.

There are two ways to transform the raw data of the treasury accounts into finished sets of yearly statistics. The traditional method is to code the information by hand, using ruled sheets and a book which contains the codes to be used in processing the data. The researcher translates the raw data directly into the numeric processing codes and enters these numbers in the appropriate fields or groups of columns on his sheets. From the latter, data cards are punched and fed into the computer.

The second or "modern" method is to use the computer to translate the raw data into the coded data to be used in processing. To accomplish this, the raw data are first translated into a storage format. Information placed in this format reads like that in the original documents except that the entry is simplified, spellings are standardized, and the various pieces of information are tagged and arranged in fixed sequences following the tags.<sup>8</sup> These changes are

<sup>7</sup> The time saved by using computer techniques will vary with the format used for storing the data (see below). Extensive coding of the data before punching on data cards or tapes will reduce the savings to the difference in time required by the machine or a human to assemble the data and manipulate the resulting statistics. Allowing the machine to sort the raw data by type and then encode it for processing should give greater savings of time, because this method shortens coding as well as assembly and manipulation time. Once the data have been placed on computer tapes in the storage format subsequent selection and coding can be done as rapidly as the machine can work, thus saving later users of the data the time required for the original translation into the storage format.

<sup>8</sup> A tag would consist of a combination of letters and numbers which serves as an identification for the field(s) of information which follow it. For example, a soldier's service record might read: REMAN (Resource-Manpower); BIOIN (Biographic Information), name, place of birth, age at time of payment, relatives; OCCIN (Occupational Information), occupation name, dates employed, rate; SERVIN (Service Information), place of service, total wages earned, discounts, net wage paid, and so on through the rest of the information in the entry. Most major computer systems supply a data storage and retrieval system

made so that the computer will be able to distinguish the various words and numbers making up each piece of information. Once all the entries have been coded in this storage format and recorded on tape, the computer is used to select those items which are of interest to the investigator. Each item is then translated into the processing codes and stored on a second tape. This tape becomes the basis for processing the data.<sup>9</sup>

The method selected by the researcher will depend in part on his knowledge of computer programming. If he is to use the second, more sophisticated method, he will need funds and archival permission either to install a keypunch in the archive or to microfilm complete runs of documents for transcription at home. Still another consideration may be his willingness to donate his data to a data bank once he has finished with them. The traditional hand-coding method lends itself to small-budget operations by persons familiar only with basic Fortran or a comparable computer language who cannot get permission to install machinery in the archive or take home long runs of microfilm. It has several disadvantages, however. Once coding has begun, any changes in the codes require the recoding from the original document of materials already finished. Also, errors in coding can be detected only by going back to the original source. Finally, the code categories will often prevent scholars with other interests from using the data cards. The traditional method thus reduces the possibility of developing a data bank which might save future scholars the same labors.

By contrast, the machine method does not initially freeze the data into any given code since the basic or raw data are stored in a machine-readable form almost identical with that of the original document. The coding for processing is done by the computer as a second step. The scholar can thus change his processing codes at will because he has the "original document" at hand. Moreover, the data he has coded for storage can be used by others for different purposes. Drawbacks to this machine method are that it is more expensive and that it may require considerable expertise in com-

which uses tags and standardized fields like those above. All are designed for use by persons without intensive training in computer programming. An example is Scientific Data Systems' "Manage." Some modification of these systems may be required for individual studies.

<sup>9</sup> For a discussion of the benefits of the "modern" format see: Laurence A. Glasco, "Computerizing the Manuscript Census," *Historical Methods Newsletter*, III (December 1969), 1-4; continued in subsequent numbers. Persons interested in these code schemes are asked to write the author at the Department of History, University of Wyoming, Laramie, Wyoming 82070.

puter programming if “packaged” programs have to be modified. Also it involves difficulties with archival regulations, not only by requiring the use of key punches or the microfilming of complete series of documents, but also because the basic data tapes are a form of publication of the original documents. Nonetheless, given cooperation by the archive and enough funds, this method promises to free future scholars from the necessity of reexamining the original treasury accounts every time they want to investigate some new problem. With all those accounts on computer tapes, the future researcher need only request (and pay for) a search of the files by the machine. In most cases this could probably be done in less time than it would take to get to the nearest airport.

Having selected a method for translating the raw data into the processing codes, the researcher then faces the problem of defining the processing codes. The rest of the essay will be concerned with this part of the proposed methodology. For the case in question, the expenditures of the Real Hacienda during the Colonial Period, the author proposes that a *Programa Para Analizar los Gastos de la Hacienda Real* (PPAGHR, pronounced “pagar”) be developed, using concepts similar to those currently employed in the formulation of Program Budgets for private businesses and agencies of the United States Government. Among the latter, that of the Department of Defense (DoD) is the best known and has served as the model for the PPAGHR.<sup>10</sup>

The program budget system is basically the application of an input-output model to the activities of a government or business. In formulating such a budget, the analysis must first order the outputs of his “firm” in terms of its major goals (Programs) and their component parts (Program Elements). Once these have been identified, the programmer draws up a list of the resources (inputs) needed to implement each Program or Program Element over a period of five or more years. Alternative ways of attaining the same goal are subjected to various types of cost and benefit analysis. Having selected a particular means, the analyst enters its current requirements in the budget. Acquisitions begin; so does a monitoring process which permits changes in the acquisition and use of all resources in ac-

<sup>10</sup> Detailed descriptions of the DoD system are: United States Government, Department of Defense, Office of the Secretary, *Study Report, Programming System for the Office of the Secretary of Defense* (Washington, D.C., 1963), and United States Government, Department of Defense, *Significant Documents On PPBS in the Department of Defense* (Washington, D.C., 1969).

cordance with circumstances such as varying costs and new policy decisions.<sup>11</sup>

In historical studies, as contrasted with contingency planning, the resource inputs and the resulting program outputs are given in the data. For the Spanish Empire in America, the acquisitions of resources (inputs) are detailed in the accounts of the royal treasurers and factors. These entries also specify the use (output) made of the goods or services, together with their cost. The historian's task is to extract from the accounts the data about these acquisitions and then to arrange them systematically as outputs in the form of yearly statistics. These statistics on royal expenditures may in turn suggest new generalizations or correlations.

The task just described can be carried out on the computer if the outputs are coded in terms of Programs and Program Elements. A Program (or Major Program, as it is sometimes called) may be defined by a general objective of the government or agency. Spanish imperial government seems to have had objectives such as Administration,<sup>12</sup> Defense, the Patronato Real, Pensions and Rewards, and possibly Public Works.<sup>13</sup> In the PPAGHR these divisions would form the basic Programs in terms of which the Crown's spending would be analyzed.

Program Elements are defined as "an integrated activity, a combination of men, equipment, and facilities which together constitute an identifiable . . . capability or . . . activity."<sup>14</sup> Each program can

<sup>11</sup> For discussions of the conceptual framework and language of Program Budgeting see: Arthur Smithies, "Conceptual Framework of the Program Budget," Chapter II of *Program Budgeting. Program Analysis and the Federal Budget*, edited by David Novick (Cambridge, 1965; second edition, 1967), 24-60. John Haldi, *Criteria for Definition of Program Categories and Elements* (Washington, D.C., 1965). George Steiner, "Program Budgeting: Business Contribution to Government Management," *Business Horizons*, VIII (Spring 1965), 43-52. Samuel M. Greenhouse, "The Planning-Programming-Budgeting System: Rationale, Language, and Idea-Relationships," *Public Administration Review*, XXVI (December 1966), 271-277. The literature on cost-benefit analysis is summarized in United States Government, Bureau of the Budget Library, *Program Analysis Techniques: A Selected Bibliography* (Washington, D.C., 1966), and the supplements to it: *Supplement I* (Washington, D.C., 1967), and *Supplement II* (Washington, D.C., 1969). A general discussion of the cost-benefit and other analytical techniques now in use is: Edward S. Quade (ed.), *Analysis for Military Decision* (Chicago, 1964).

<sup>12</sup> Administration would contain subdivisions such as Justice, Treasury, Police, Political Administrators, etc.

<sup>13</sup> Pensions and Rewards and Public Works might be consolidated under one heading such as "Social Investment."

<sup>14</sup> U. S. Government, Department of Defense, *Study Report*, II-1.

be classified at various levels into these component parts. In the PPAGHR each Program may be broken down into four or five stages.

These definitions of Programs and Program Elements, together with an examination of the Program Element List prepared for the Department of Defense, show that each element contains a limited explicit specification of the resources needed for that capability or activity. Reading the records used in the Department of Defense system, one would find a Program Element Code, followed by code fields for the Resource Codes, which are divided according to the manpower or service requirements of the element, its "forces" or materials requirements, and its costs. Thus each Program Element has attached to it a file listing the resources needed. The total file of Program Elements or outputs is in effect a relisting of the total resource inputs. The file serves as a bridge between input and output categories. Because of high speed data processing, any part of the resource file (inputs) can be reconstructed from this output file.

For historical purposes, the basic codings of the resource data by Programs and Program Elements would have to be supplemented by a code for the date of payment for the good or service,<sup>15</sup> a geographic code indicating where the items were used, and a code which would specify the purchasing treasury. This last code could be combined with an account serial number to specify the sources from which the data were taken.<sup>16</sup> Depending on the accounts used some or all of this additional information would have to be added to the raw data, regardless of the method used to translate the data into the processing codes. The processing format being used by the author arranges these code fields as follows (left to right): date, treasury and account serial numbers, PPAGHR code, geographic code, resource codes—manpower (with quantity), materials (with quantity), and cost—and then all other codes.

<sup>15</sup> Most entries in the accounts contain several dates. The author's research has concentrated on studying the actual yearly cost of defense. Consequently, the preferred date is the date of the *carta de pago*. If that is not given, then the date of *libranza* is preferred. Lacking that, a date supplied from other sources is preferred over the date the entry is entered in discharge (*data*) or account (*en cuenta*). Dates of entry into *data* or account are the least desirable because expenditures sometimes are not entered into the accounts until years after the funds were paid out. See, for example, the expenditures for the outfitting of the *San Pelayo* for the conquest of Florida in 1565 which are entered in the treasury accounts of the Casa de Contratación for 1578. AGI:CD 310B, No. 2, No. 12, 431:2-473:1.

<sup>16</sup> Archival citations are too complex for use in machine processing. Each treasurer's and factor's accounts would need to be identified (even those which are no longer extant) and given a serial number.



Once the data have been coded in this processing format, a relatively simple computer program is used to convert the cost into a constant unit of account (maravedí) and then to sort the expenditure data according to the PPAGHR and geographic codes and the year in which payment was made. These statistics are then printed out as tables by means of a magnetic tape. Using the figures stored on the tape, further statistical manipulations can be performed to obtain such things as five-year moving averages, least squares slope values, etc.

On the other side of the ledger—that showing the income of the Crown—an approach similar to the one described here might be developed. Revenues could be classified by type and then subjected to machine analysis. For reasons already indicated, the resulting statistics would have less reliability than those generated by studies of royal expenditures. If the data are coded first in the “modern” storage format discussed above, the income entries might be used to make studies of population on the basis of tribute payments, port traffic, and the composition of the merchant and other classes, to name but a few uses. Some work along these lines has already been done, but usually without reference to its integration into a larger data set which could be reexamined by others, from the “raw” data up.<sup>17</sup>

Several obstacles lie between this proposal and the use of a computerized methodology to produce series of statistics on various aspects of colonial life and on the Crown’s spending of its colonial revenues. First, interested scholars need to work out and publish the formats and tags for storing the data in the “modern” non-numerically coded form. Given this information, individual scholars might prepare their data so that they could be integrated into a data bank. Second, those interested in analyzing royal spending in terms of Programs and Program Elements would need to complete the PPAGHR coding system. Third, individual projects would be needed to perfect computer programs and generate the data to be placed in the bank. Finally, if suitable institutional arrangements could be made, a data bank might be created which would make the treasury data machine-retrievable through the use of tested programs. This bank would make it possible for future historians of the colonial period to use these quantifiable data for different purposes and without having to retrace the steps of others in the sources.

In the end, our understanding of the Colonial period will be enriched not only by the new knowledge brought to light by the quan-

<sup>17</sup> Rodríguez Vicente, “La Contabilidad,” *passim*.

titative studies which produced the data for the bank, but also by the wider perspective which an individual scholar can gain for the same expense in manhours of research time. The computers for these studies are at hand on university campuses throughout the country. The question would seem to be whether interested Latin American historians can cooperatively perfect the codes and formats required and build a data bank for future use.