Volume Title: The Role of the Computer in Economic and Social Research in Latin America

Volume Author/Editor: Nancy D. Ruggles, ed.

Volume Publisher: NBER

Volume ISBN: 0-87014-260-7

Volume URL: http://www.nber.org/books/rugg75-1

Publication Date: 1975

Chapter Title: The Computer in International Price Comparisons

Chapter Author: Zoltan Kenessey

Chapter URL: http://www.nber.org/chapters/c3779

Chapter pages in book: (p. 313 - 318)

THE COMPUTER IN INTERNATIONAL PRICE COMPARISONS*

ZOLTAN KENESSEY

United Nations, New York

INTRODUCTION

The use of the computer in handling large price files, the subject chosen for his paper by Jorge Salazar-Carrillo, is one of the new examples of the use of the computer in economic and social research. The use of the computer in this area—more specifically in the simultaneous handling of price files for a number of countries and with a view to computerizing a great deal of the work of comparing such prices—is rather novel not only in this region of the world, i.e. in Latin America. One has to remember that the pioneering effort in this field of research, the study undertaken in the Organization for Economic Cooperation and Development (OECD) by Milton Gilbert and Irving Kravis in the early 1950s, could not yet rely on the computer handling of large price files. Neither was the computer involved in price-data operations required for the comparison of the socialist countries performed in the late 50s and early 60s in the Council for Mutual Economic Assistance (CMEA). The earlier Economic Commission for Latin America (ECLA) study in this area was not computerized either. Indeed, there is little previous experience in the field; and among the few exceptions, the experience of the Ruggleses in the management and computer manipulations of the price data for Latin American countries should be mentioned.

The current situation is very different in this respect. All the efforts at international price comparisons carried out on any significant scale are now computer-oriented. The UN International Comparison Project, a cooperative undertaking carried out by the UN Statistical Office and the University of Pennsylvania, relies heavily on the computers in New York and Philadelphia in handling price data (and some other tasks involved in the comparisons). The comparisons of the Statistical Office of the European Economic Communities, which are undertaken in a coordinated manner with the UN International Comparison Project, are also computer-oriented. And, as one can see from the paper presented by Salazar-Carrillo (as well as that by Howe and Villaveces) to the present meeting, the comparisons concerning the countries of Latin America carried out by the Brookings Institution in cooperation with a number of research institutions of Latin America are also heavily involved in computer operations.

Before going into the details of the discussion it appeared helpful to remember that the use of the computer for this kind of international comparison is a novel feature in economic and social research. Following this introductory thought, the more detailed remarks on the handling of large price data files by computers is offered in the following order: the merits of the study; findings on the positive side with the use of the computer; negative experiences with the use of the computer; and a few remarks.

* Comments on the paper of Jorge Salazar-Carrillo appearing in this volume. The views expressed in this paper are those of the author and do not necessarily coincide with those of the United Nations.

1. The Merits of the Study and its Positive Findings

The basic purpose of the present conference is well served by the report of Jorge Salazar-Carrillo, which can be considered as a very helpful, honest, and frank paper on some main issues of the use of the computer in this field. The paper does not try to becloud the issues; it depicts both the positive and the negative as experienced by Salazar-Carrillo and his colleagues at Brookings while carrying out this particular piece of research. Obviously some of the experience reported is subjectively based: nonetheless, it is both valuable and relevant to our understanding of the problems related to the use of computers, since such problems, in fact, can be both objective and subjective in nature.

In keeping with the objectives of the paper and the conference, one should refrain from a discussion of the economic assumptions or statistical methods of the investigation and should avoid reviewing the ECIEL [Program of Joint Studies on Economic Integration in Latin America] study in general. In this respect, one should only stress that everybody concerned should be pleased with, and indeed grateful for, the efforts of the Brookings Institution in this important field of investigation, carried out in cooperation with a large number of Latin American research institutions of the stature of the Fundación de Investigaciones Economicas Latinoamericanos in Argentina, the Instituto Brasileiro de Economia of the Fundaçao Getúlio Vargas, the Centro de Estudios sobre Desarrollo Económico (CEDE) of the Universidad de Los Andes and the Departamento Administrativo Nacional de Estadistica (DANE) in Colombia, and the Centro de Estudios Economicos and Demographicos of El Colegio Mexico. It appears of wider international significance that a sustained research and training effort of this kind became a reality in the Western Hemisphere.

Leaving aside the relative merits of the economic assumptions and the statistical methods chosen for the ECIEL study and the various methodological decisions taken in guiding the work (for example, the views described by Salazar-Carrillo in part 2.7 of his paper concerning "Analysing Intercountry Price Dispersions") and concentrating on the computer problems described by the author, the following remarks are in order.

The exposition of the major sequences of the data operations and related problems is generally quite clear. Figure 1 of the paper (p. 296) is especially helpful for understanding the various stages of the work and identifying the place of the computer operations in them.

The summary of the experiences gained is also well designed, particularly concerning the definitive pattern observed in the use of the computer in the ECIEL study. At the early stages of the study, when the data were being processed for organizational and storage purposes, heavy reliance was put on the computer. The degree of computerization was less intense in the cleaning and editing stages. During the final estimating stages, virtually all the work was done by machine.

It is one of the merits of the paper that it does not avoid dealing with the evaluation of the day-to-day, down-to-earth operations. As an example of this, reference can be made to the review given by Salazar-Carrillo on the possibility of using the computer to check the coded results of the price surveys, the original questionnaires, and the standard checking forms against each other for possible errors and dis-

crepancies. On the basis of this review, it was decided not to use the computer for this task but to rely on research assistants aided only by desk calculators. The reasons for this decision were various, among them the bottlenecks in keypunching at the Brookings Electric Data Processing unit at the time.

In respect to the general findings of the study, it has to be underlined that Salazar-Carrillo's overall evaluation of the use of the computer is positive. He especially emphasizes those advantages of the computers in the handling of large price-data files which permit the computation of cross-country indexes on the bases of various index number formulas at reasonable cost.

2. Negative Experiences with the Use of the Computer

In many ways, the negative experiences of the ECIEL study are more interesting than the positive ones. Therefore, one should dwell on them somewhat longer. However, an attempt is made in these remarks to attach a somewhat different interpretation to a number of the negative findings than is offered by the author.

One of the negative experiences of the study is related to the test for detecting outliers in checking the dispersion of prices within each country. In this work, the assumption accepted by Salazar-Carrillo was that "unreasonable price averages can result from one or more extreme observations in the sample prices of a particular product." (In parentheses, one is tempted to observe that neither the researcher nor the computer is in very good shape when it comes to deciding what price averages are "unreasonable" ones.) In order to screen the outlying extreme values, the means and the standard deviations were used. For various reasons, this test did not prove to be tremendously helpful, and finally it was decided not to involve the computer in this operation.

There is no apparent reason to question the wisdom of this decision. It is quite possible that under the given circumstances, a noncomputerized procedure was actually preferable. But it is certainly less obvious whether the computer or the "circumstances" should be blamed for the difficulties.

The test was based on the relationship of the mean and the standard deviation, following certain assumptions concerning the distribution function of the universe and the size of the samples involved. With the estimation of confidence intervals it was hoped to arrive at "reasonable" limits, outside of which there was a low probability that a value observed would belong to the same universe as that from which the sample was generated. Obviously, the successful application of the test chosen depended on a number of factors, among them on the adequacy of the size of the sample. Unfortunately, the size of the sample compared to the universe of prices was small. So the question posed by the author in one of the subtitles of the paper, namely, "Objective outlier tests or subjective evaluation procedures?" (p. 300), loses very much of its apparent generality and relevance for judging the relative efficiency of computer operations for such purposes. Under improper conditions, the use of the computer will obviously be less than helpful both for such and many other operations.

Naturally Salazar-Carrillo is aware of the general requirements of good price samples. However, in part because of the pilot character of the project and due to certain—under the circumstances probably unavoidable—limiting factors, the

price samples could not be very close to the ideal. The data obtained in the ECIEL price sample referred to particular segments of the reference year in time and to particular geographical segments—capital cities—of the given countries. Compared to the whole universe of prices (both in the dimensions of a whole year and with respect to the prices all over the countries and not only in their capital cities) the selection of prices was a relatively small sample of all the prices in any given country in the year of observation.

Therefore, while one need not question at all the wisdom of the particular decision in the ECIEL project to avoid the use of the computer in testing the irregularities in the price data, one can still doubt the validity of any implied or explicit generalization concerning the usefulness of the computer for such studies carried out under different circumstances.

It is interesting to note that on the basis of the ECIEL project, Salazar-Carrillo considers that once the price data are clean, the computer is essential to the calculation phase. Again, one should like to agree with him. Certainly after the data problems of an external character are eliminated, the efficiency of the computer operations should become much greater. Computers are not particularly helpful in eliminating external data problems and obviously, to avoid the proverbial "garbage in—garbage out" situation with the computer runs, a very considerable effort was required in the ECIEL project to improve the quality of the basic data before further manipulation was considered feasible. In a sense, the need for such efforts is completely unrelated to the computer, since whatever method of computation is chosen, reliable basic data are essential inputs for them. Whether manual computations or machine runs follow, this type of effort is indispensable. Unless one is willing to regret the discipline required by computer runs in data preparations and implicitly considers the computer as the "cause" of such efforts, some of the remarks in the paper are not to the point. In this connection, one may find problematic those references in the paper which voice complaints about the time spent on the computer applications in the project. Naturally, this is not to question at all the facts described or the subjective feelings of those involved. It is in the interpretation of these facts that a different view is offered.

For example, the paper points out that in terms of costs and efficiency "there were for all practical purposes, no alternatives to the computer in the calculations phase, even though some of its general disadvantages were still present" (p. 308). The main disadvantage appears to be that "the additional burden of learning at least the basics of computer science and managing the computer aspects of research projects falls upon the researcher" (p. 292).

While this may be quite true, it hardly justifies classifying the price thus paid in the research as a "disadvantage." To use an analogy, if one wants to operate an automobile, it is unavoidable to learn to drive. This is a process which takes time, usually there is some amount of frustration or anxiety involved in it, and it even costs money. Yet, it would not be very illuminating to say that compared to walking, driving had this particular "disadvantage." (Also while we don't remember it—learning to walk is not that easy either.)

At an earlier session of the conference, Sadowsky referred to the larger gap that exists in the social sciences between computer specialists and the researchers than that observed in the natural sciences. Part of his explanation of this gap was

given in a reference to the greater tendency of the social sciences to be what he called "one man shows," as compared to frequent team performances in the natural sciences.

It would certainly not be justifiable at all to classify the ECIEL project as a "one man show"—and this is not to lessen the significance of the contribution of such key individuals in the project as Joseph Grunwald and Jorge Salazar-Carrillo. However the ECIEL project (which is really a group of projects beyond and over the price-data operations now being discussed) is clearly a very complex, very ambitious undertaking, and the resources available for it are not quite proportionate to the width and depth of the task.

The tasks of the ECIEL project are of such a complexity from an operational point of view that the limitations of private research institutions—even if allied with each other in a cooperative manner—are quite severe. This is not only a matter of the limitation of the resources available at the center of the ECIEL project. According to the experience gained in the UN International Comparison Project, a very crucial role has to be played by the national statistical offices in the development and carrying out of such projects. Almost invariably, such offices are in a much better position to carry out the heavy statistical burden involved in the work than even well-organized private research groups. (Howard Howe and Roberto Villaveces give some excellent insight into these problems in their extremely useful paper.)

3. Concluding Remarks

Depending on the task, the computer can be a reliable tool and ally of the single researcher, the research group, and the large institution. But obviously it can fail to serve any of them—at least, for certain tasks and under certain circumstances.

It is very likely that the efficient use of the computer in the handling of large price-data files may require a somewhat larger-scale operation than was possible in the ECIEL framework. But irrespective of whether this is the case in this particular instance, it may be helpful to recall certain simple economic relationships which are taught to all students in other contexts, but which are often not in the forefront of the thinking when consideration is given to the use of the computer in economic research.

Just as in everyday economic activities certain proportions have to be maintained in the combination of the various factors of production, the work in economic research should be considered subject to requirements of a similar nature. It would appear, for example, that the use of the computer in research is not dissimilar to capital inputs in production and that the combination of computer use with the other types of inputs (like the time of the researchers) involves a great deal of managerial-type decisions.

In any given research, obviously not every combination of the capital, labor, and management inputs will be optimal or even feasible. It is surprising that the use of the computer in economic and social research is rather seldom analyzed from this angle. The usual approach to the study of the global economics of the use of the computer in economic research is much simpler, and in fact it is often

restricted to the comparative costs of computer operations and manual ones. Consequently, our knowledge of the economics of the use of the computers in economic research is less satisfactory than one would expect in the case of a profession dedicated to the optimal allocation and use of resources.

Therefore, one should be grateful that the National Bureau of Economic Research and its sister organizations in Mexico, Brazil, and Argentina found it possible to organize a conference to highlight these problems.