THE USE OF THE COMPUTER IN HANDLING LARGE PRICE FILES: THE EXPERIENCE WITH A BENCHMARK COLLECTION IN LATIN AMERICA

JORGE SALAZAR-CARRILLO*
The Brookings Institution

SECTION 1

1. *Introduction*

Although the influence of the computer has permeated many fields of study, its use in price research has been rather limited, especially in international comparisons. This may be due to the nature of empirical work in this area; specifically, the number of observations handled and variables involved, which place such efforts between time series and surveys in terms of size and complexity.

Would a larger role for the computer in this field be beneficial? Could its present contributions be increased and how? As general background in answering such questions, an outline of some of the seemingly positive and negative aspects of the computer, as far as they are applicable to research projects in the area of prices, would appear worthwhile.

1.1 *Positive Aspects*

One of the potential advantages of computer applications lies in the quickness it allows in the processing, organizing, and cleaning of the data. This reduces the delay between the actual collection and the presentation of the final results.

Data processing and cleaning would appear to respond well to machine operations not only in terms of time, but also in the form of other cost reductions. Having the information readily available in well-organized tape or disk files, plus checking and editing by computer, appear to be the pillars of this greater efficiency.

Another apparently important factor on the credit side is increased accuracy in research operations. This advantage applies to all the stages of a project: from the cleaning and checking phase to the final manipulation of the data.[1]

An additional benefit is that computer operations necessitate that the data be arranged in a highly organized and explicit form, forcing the creation of files, banks, and other orderly information structures. This facilitates data retrieval, while encouraging efficient filing, which is central to accessibility of data to other researchers.

* The author is senior fellow with the Brookings Institution. He wishes to thank Joseph Grunwald for his continued encouragement, and Ricardo Martinez and George Plinio Montalván for general help. The author is also thankful to Philip Musgrove for his valuable comments, and to Stanley Braithwaite, Robert Ferber, Irving Kravis, and Richard Ruggles for advice on some aspects of the data processing operations discussed here. The views expressed in the paper are the author's and do not necessarily reflect those of the staff, officers, or trustees of the Brookings Institution. The research reported here has been mostly financed by Ford Foundation grants.

[1] This is partly compensated by the fact that some errors can be introduced in the computer handling of the data, as in keypunching and card reading.

The importance of rapid retrieval for research and experimentation is obvious. The availability of data files which are either easily transferable or accessible through computer linkups is also of great importance. This allows other researchers to use the data arising from a particular research project for other research purposes or for an alternative formulation of the same research goals.

The computer has no equal in its ability to undertake complicated and long manipulations of data. Computers permit experimentation with several alternatives, while previously cost considerations limited it to only two or three. For example, for large price files, the computer would permit the testing of the effects of different index formulas on the measurement of price change, at reasonable cost.

This characteristic, together with the fact that it provides rapid feedback, makes the computer quite amenable to hypothesis testing. Techniques which have relied heavily on these computer properties are those of economic simulation and sensitivity analysis.

1.2 *Negative Aspects*

On the other hand, computer applications could create problems. First of all, in the present state of computer technology, the additional burden of learning at least the basics of computer science and managing the computer aspects of research projects falls upon the researcher. For large projects, this task may be placed in the hands of specialists. However, these specialists must still be understood and controlled by the economist or statistician. This necessity may result in some waste in the use of human resources, as the attention of the researcher has to focus on matters in which his productivity is lower.

Second, the computer appears to diminish the independence of researchers, as they now have to rely heavily on the efficiency of their computer center. The alternative of buying computer services outside is sometimes open. However, just considering this possibility is rather time consuming, while again forcing the researcher into acquiring special skills not of his immediate interest.

Basically, as a result of communication problems with programmers, a large amount of research time has to be devoted to carefully checking computer output. In a relatively large number of cases, computer exercises end up being repeated. Facing these conditions, economists and statisticians frequently decide to learn programming, and end up by becoming their own programmers, a solution which negates specialization.[2]

SECTION 2

2. *Computer Operations and Price Research: An Illustration*

In the previous section, the advantages and disadvantages of the computer for price research have been discussed at a general level. It is now time to turn to a specific consideration of the potential use of computers in this area, and to

[2] Most of the difficulties introduced by the computer involve deficient communications. One of the most befuddling characteristics of computer operations is the extent to which its specialists persist in speaking "computerese," a complicated and confusing language extremely specific to their problems. Although every discipline is more or less affected by this problem, computer people seem to be more oblivious of the confusion this creates than are other specialists. This situation is particularly shocking because computer personnel basically serve other researchers, their output being more characteristically intermediate than those of other disciplines.

an examination of the ways in which its net contribution could be enlarged. This will be done first by way of a practical illustration, using a large benchmark survey of prices undertaken in Latin America in 1968.[3]

2.1 *Purpose, Nature, and Scope of the ECIEL Study*

It is necessary to enter a little bit into the purpose, nature, and scope of the ECIEL study of which the survey is part. For further reference two previous papers by the author should be consulted.[4]

The main purposes of the study are: (a) to construct an international price index for Latin America; (b) to determine purchasing-power parity rates for these countries; and (c) to estimate their gross domestic products in real terms. A secondary objective is the study of price and cost structures from the points of view of efficiency and trade (especially that among Latin American countries). Hopefully, such research will improve the understanding of integration and development in the area.

For the preparation of the basic estimates, the study required the collection of price and expenditure information. These were obtained in great depth and detail for 1968. The gathering of these data took place in all the countries involved under a common methodology and common procedures.

These estimates cover different categories of goods and services. They correspond to the main subdivisions of the gross domestic product, as calculated from the expenditure side. However, this paper will concentrate on the private-consumption sector, and particularly on the collection of private-consumer-goods prices.[5]

[3] Brookings undertook such research in collaboration with thirteen research institutions from Latin America. All these institutions participate in a unique research venture coordinated by Brookings under the acronym of ECIEL. ECIEL stands for the Spanish equivalent of Program of Joint Studies on Latin American Economic Integration. About twenty research institutions in Latin America participate in its several projects. For a fuller description of what ECIEL is, see Joseph Grunwald and Jorge Salazar-Carrillo, "Economic Integration, Rate Exchange and Value Comparisons in Latin America," in *International Comparisons of Prices and Output*, D. J. Daly, ed., New York, NBER, 1972. The ECIEL institutions participating in this particular study are: Fundación de Investigaciones Económicas Latinoamericanas (FIEL), from Argentina; Instituto de Investigaciones Económicas of the Universidad Mayor de San Andres and Instituto Nacional de Estadistica, from Bolivia; Instituto Brasileiro de Economía of the Fundaçao Getúlio Vargas, from Brazil; Instituto de Economía y Planificación of the Universidad de Chile, from Chile; Centro de Estudios sobre Desarrollo Económico (CEDE) of the Universidad de los Andes, from Colombia; Instituto Nacional de Estadistica, from Ecuador; Centro de Estudios Económicos y Demográficos of El Colegio de México, from Mexico; Centro Paraguayo de Estudios de Desarrollo Económico y Social (CEPADES), from Paraguay; Centro de Investigaciones Sociales, Económicas, Políticas y Antropológicas (CISEPA) of the Universidad Católica del Perú, from Peru; Instituto de Estadística of the Universidad de la República del Uruguay, from Uruguay; Banco Central de Venezuela and Centro de Desarrollo (CENDES) of the Universidad Central de Venezuela, from Venezuela.

[4] See Joseph Grunwald and Jorge Salazar-Carrillo, *op. cit.*, and Jorge Salazar-Carrillo, "Price, Purchasing Power and Real Product Comparisons in Latin America," *Review of Income and Wealth*, Series 19, No. 1 (March 1973), pp. 117–132.

[5] The reasons for choosing these goods for the illustration attempted here are:

(1) It is by far the most important sector, covering about 75 percent of gross domestic expenditures in the countries in question.

(2) The size and complexity of the information gathered in this sector makes it more amenable to computer handling and thus better for illustrative purposes. The number of observations per item, as well as the number of items, are much higher in the private consumption sector. Moreover, each price is identified as to quality, type of store in which it was collected, and the income stratum of the shoppers using the store. Not as much information was sought in connection with the rest of the prices in the survey.

There were 416 items included in the price sample for private consumption. Three qualities were priced in each item, adding up to 1,248 individual products. As an average, about seven or eight price observations were to be collected for each quality.[6] Thus, the maximum numbers of observations a country was expected to report came to about 9,000. Yet, a much smaller number were actually reported in the end, due to unavailabilities, price controls, poor fieldwork, and so on.

The level of disaggregation attempted was quite ambitious. For example, private consumption was ultimately divided into more than 100 subcategories. As to geographic coverage, it is important to note that not all the Latin American countries were included, the collection being restricted to the LAFTA (Latin American Free Trade Association) countries.[7]

Up to now, the work on prices has concentrated on the most important city in each nation.[8] However, the relationships will be adjusted to correspond to the nation as a whole by way of the regional price-index information available in the various countries.

2.2 *The Preparation, Execution, and Coding of the Benchmark*

The survey was prepared carefully; many months of work with the price indexes, national accounts, expenditure surveys, and so on, of the LAFTA countries were put in before the testing of alternative LAFTA baskets began. After a common basket was chosen and specified, the uniform methodology and procedures were set down. The final phase before the execution of the survey called for some testing, and for the training of supervisors and interviewers. All this was basically coordinated by Brookings.

A standard questionnaire was used in all countries. In fact, the price collection took place during the same month (May, 1968). To insure comparability in the goods selected for pricing, these were carefully specified. The same specifications were used in every country. Furthermore, a coordinating commission traveled to all countries in order to select (with the help of each collaborating institution), among all those varieties, brands, and models that complied with the specifications, those that were comparable in the different countries.

After the survey was completed, the data were given a thorough checking by the field institutes. When the corrections suggested by these verifications were done, the information was coded by the institutes. For this, the uniform coding sheets prepared at Brookings were used. Such a coding approach provides the first illustration of the application of computer operations in the price survey.

Although the Brookings Computer Center had been in existence only a short time by early 1968 and was, therefore, still facing some organizational problems, the decision was made to try to benefit as much as possible from its potential contribution. Thus, it was felt that in order to record and organize the data to be sent by the institutes as efficiently and cheaply as possible, it was convenient to homogenize the whole operation by distributing standard coding sheets to all institutes. These would then be filled out and sent in by every collaborating institute.

[6] The number differed among the major categories.
[7] See footnote three for a list of countries.
[8] With the exception of Ecuador, where both Quito and Guayaquil were covered.

Most importantly, these coding sheets were to be designed so as to streamline their keypunching.

This solution seemed to have various advantages. First, it removed the inefficiency involved in having to work with diverse types of coding sheets. Second, it speeded up keypunching, as well as making it less prone to error.[9] Third, having the data on cards made it easier to reap the potential benefits of certain computer operations.[10] Fourth, homogeneous data listings could be generated with ease, facilitating the verification and cleaning of the survey at Brookings.

On the other hand, some problems could be foreseen in adopting this approach. Misunderstandings could be expected in the field, as a result of which the coding sheets might not be filled out properly. Also, for certain institutes, country-specific procedures would have been preferable for several reasons.

In the end, the experience with this solution was somewhat mixed. Some institutes, because of various problems, failed to send the uniform coding sheets or did not fill them out properly. Then, inconsistencies were found between the information coded and that included in the questionnaires. Still, for the great majority of the institutes, no major problems were encountered, and the use of the coding sheets appears to have realized its expected gains in efficiency. The computerization of the study right from the beginning appears to have been beneficial for the study, in general, and for the initial phases of registering and checking the information, in particular.

Thus, in price studies across countries or regions, the uniformity of data-sources (principally coding sheets, but also questionnaires and field verification forms) seems worthwhile. However, it might be necessary to spend considerable time in examining the particular situation of every country or region before preparing these sources. Then, explaining the standard forms to supervisors, interviewers, verifiers, and coders, as well as training them in their use, would probably be essential.

2.3 *The Revision of the Coding Sheets*

The questionnaires used in the price collection and the standard forms utilized in checking the fieldwork were also sent to Brookings by the institutes.[11] The need for checking the coding work done by the institutes was firmly established by a cursory examination of these forms, in conjunction with the coding sheets. This set the stage for the next step in the data-handling process (see Figure 1).[12]

This step provided another opportunity for the application of computer methods. Given that the three data sources contained the same basic information, it was possible to establish a correspondence among them without much difficulty.

[9] It should be made clear that unlike most empirical research efforts in the area of prices, in this project an attempt was made to record and save the original price observations.

[10] To wit: the processing of the data could be expected to be fast; its organization explicit and well defined; and the information it conveyed ready for varied types of statistical analysis.

[11] Copies of these materials were returned to the institutes in case further field verification became necessary.

[12] At this point, it may be convenient for the reader to go over the various data-handling operations performed in the study, which are shown in Figure 1. Many of them will be discussed later on. Still, a review of such a chart may help the reader grasp the direction and objectives of the processing, cleaning, and editing operation performed, as well as their interconnections.

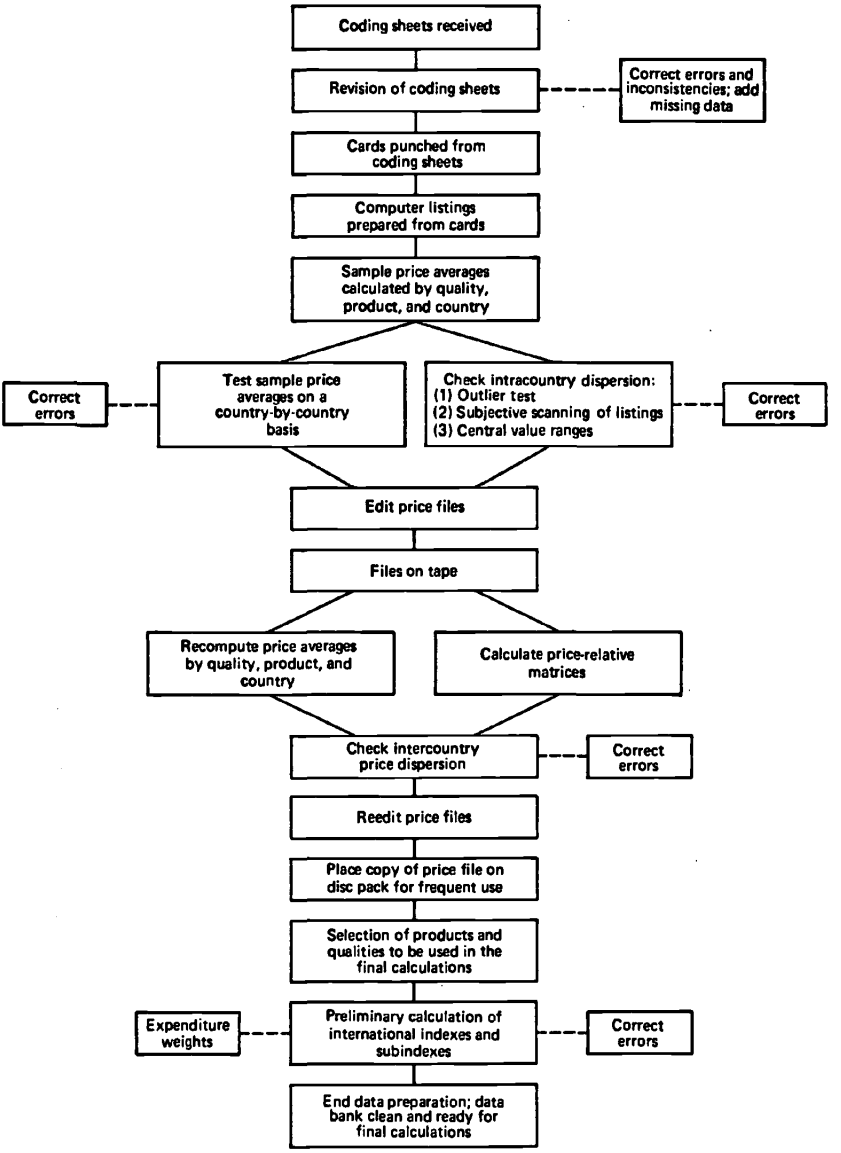

FIGURE 1 DATA HANDLING OPERATIONS IN THE ECIEL PRICE SURVEY

Using such correspondence, the data from each source could be keypunched and a program written to spot those cases in which they did not coincide. A review of the different data-sets could establish those cases in which the values in the coding sheets should be considered questionable.[13] If considered appropriate, the coding sheets would be edited at that point; if not, additional field verifications would be performed.

However, in this case, consideration was given to alternative ways of determining if such discrepancies existed. In particular, the possibility of checking the presence of such inconsistencies by hand was thought to be appealing. After close study, this technique was deemed superior because:

1. The keypunching of the three data-sets was found to be quite costly. This was partly determined by the facts that each data-source was organized somewhat differently, and that even for the same form (questionnaires, coding sheets, field verification forms) there were dissimilarities among countries.

2. At the time such checks were required, keypunching involved serious delays at the Brookings Computer Center.

3. With the use of the traditional procedures, the two steps involved in the revision of the coding sheets (scanning for discrepancies and determining their causes by checking the various sources) could be combined into one.[14]

2.4 *Checking Price Averages Within Each Country*

After the inconsistencies in the data had been corrected and other errors and omissions in the coding sheets were taken care of, the information included in them was punched into cards. If field verification was still pending in connection with the revision of the coding sheets, provisions were made to incorporate into cards any possible corrections or inclusions as soon as they came in. From these cards, computer listings were prepared for further checking. These listings showed the country, income stratum, type of store, and quality level corresponding to each particular price observation (see Table 1 for an illustration).[15]

The next step in the data-handling system called for average prices for each product and quality to be calculated. These averages were then contrasted with two independent sources: the average prices extrapolated from the 1960–1962 ECLA (Economic Commission for Latin America) survey and the average prices used in the calculation of consumer price indexes in the cities surveyed.[16]

A correspondence was first established between the items and qualities in the ECIEL [Joint Studies on Latin American Economic Integration] study and those in the ECLA survey and in the various consumer-price-index baskets in LAFTA. For those items and qualities found to be similar, the corresponding prices were compared. Whenever wide discrepancies occurred, the products in question were scheduled for closer examination. The specifications of the items in the different surveys were checked again. Those whose definitions were found to be somewhat

[13] The information contained in the coding sheets was considered the most credible, because these were supposed to be filled in after the preliminary field checking had been completed.

[14] The second step had to be done by hand anyhow.

[15] These listings were regularly updated as new data came in.

[16] See U.N., Economic Commission for Latin America, *A Measurement of Price Levels and Purchasing Power of Currencies in Latin America, 1960–62*. E/CN.12/653. 1963.

TABLE 1
EXAMPLE OF COMPUTER LISTINGS GENERATED FROM CARD DECKS, 1968 ECIEL BENCHMARK, LAFTA COUNTRIES

Country Code	Type of Store Code	Income-Strata Code	Product Code	Superior Quality	Intermediate Quality	Inferior Quality	Questionnaire Number
01	1	03	1111013	380.00	–	280.00	3
01	1	01	1111013	410.00	320.00	280.00	1
01	1	03	1111013	350.00	260.00	220.00	2
01	2	03	1111013	320.00	320.00	260.00	4
01	2	03	1111013	–	260.00	270.00	5
01	3	01	1111013	390.00	350.00	320.00	6
01	3	03	1111013	450.00	350.00	280.00	7
01	3	03	1111013	450.00	320.00	300.00	8
01	1	03	1111013	240.00	200.00	200.00	3
01	2	03	1111013	260.00	240.00	160.00	4

diverse were not checked further. For the rest, the ECIEL averages were verified by referring back to the original sources and by requesting additional field research when necessary. In the end, the card decks were updated to reflect the changes introduced in this process.

Most of this checking could have been undertaken by computer; however, traditional hand-calculation methods were preferred. Nonetheless, the computer was used as an aid in some of the tasks involved. The information conveniently printed and summarized in the computer listings, as well as the computer-calculated price averages, was central to the hand methods used.

The characteristics of the tests conspired against a more intensive use of the computer in this case. Before the average prices could be compared to each other, their similarity had to be established. Computer operations were not amenable to this kind of work, as an evaluation of the compatibility of the various specifications was required.

Furthermore, only the ECIEL price information was organized in a fashion that would make the application of computers straightforward. The consumer-price-index information and the ECLA extrapolations were provided in work sheets. Coding and punching these into cards would have been required before computers could be used.

Then, even though products with dissimilar specifications were to be screened out as a first step, some differences were bound to remain which could validly cause price discrepancies. Thus, any machine test comparing average prices from different surveys would be partially inconclusive anyhow and would need to be supplemented by subjective evaluation.

2.5 *Checking Price Dispersion Within Each Country*

2.5.1 Outlier tests. Unreasonable price averages can result from the influence of one or more extreme sample prices. It is important, hence, to check for price variation at the same time that the sample price averages are being questioned. Whenever such troublesome observations are detected, they could, depending on their causes, either be corrected, eliminated, or left standing. Some of the suspect price averages may be straightened out in this process.

Such dispersion checks were felt to be machine adaptable and a program was written to implement them. This program computes and prints the mean price for each product and country, as well as its standard deviation (see Table 2 for an example).

TABLE 2
SAMPLE PRINTOUT OF OUTLIER PROGRAM. 1968 ECIEL BENCHMARK (C = 2)[1]

Product Code	Mean (M)	Standard Deviation (S)	Lowest Value	Highest Value	M − C(S)	M + C(S)	Outliers
1111013	3.77	0.63	2.70	4.50	2.52	5.03	none
1111013	3.07	0.37	2.60	3.50	2.34	3.81	none
1111013	2.61	0.45	1.80	3.20	1.72	3.51	none
1111023	2.41	0.32	1.90	2.80	1.78	3.04	none
1111023	2.26	0.38	1.80	2.80	1.50	3.03	none
1111023	2.05	0.33	1.60	2.70	1.39	2.71	none

[1] C stands for an arbitrarily selected number used to define. together with the standard deviation (S). an interval about the mean (M).

In order to screen the outliers, or extreme values, the relationship between the mean and the standard deviation was used. The key question is how many standard deviations removed from the mean must a value be for it to be considered suspect? The rule to follow depends on the assumptions that can be made about the distribution function of the universe, as well as the size of the samples involved. When a normal or a *t* distribution can be assumed, it is known that the sample mean and standard deviation can be utilized to estimate confidence intervals, outside of which there is a low probability that the values found would belong to the same universe, or population, from which the rest of the sample is taken. Still, the selection of limits outside of which observations should be deemed suspicious is usually an arbitrary and subjective exercise.

In order to determine outlier values as *objectively* as possible, a procedure was used in which several confidence intervals were tested and, depending on statistical criteria, a particular one was chosen as the most appropriate. Intervals ranging from one to three standard deviations, at both sides of the mean, were experimented with. For this purpose, the program calculated the confidence intervals about the mean for a varied number of standard deviations. These intervals were printed out explicitly (see Table 2).

The results of the experimentation showed that confidence intervals between 1.5 and 2 standard deviations were the most reasonable from an objective standpoint. The program was then run again with the chosen values, with the prices falling outside the predetermined confidence intervals being printed out. These prices were then submitted to further tests in order to decide if they should be deleted from the sample.

Although operationally this test and the program that embodied it were quite successful, doubts about its practical and theoretical appropriateness were raised. To discuss them, it is important to remind the reader that three distinct qualities were surveyed for each of the items included in the survey, with at the most nine prices collected per quality.

Given that each quality could be considered as a separate product, the test was made to center on the dispersion of the price observations around the mean for *each quality*. When the outlier test was put to work, it was found that when observations fell outside the confidence intervals, there were usually one or, at most, two of them. Moreover, a superficial examination of the computer listings seemed to indicate that the observations spotted were not necessarily erroneous, while on the other hand, many irregularities had been missed by the test.[17]

The fact that commonly one, and rarely two, observations were flagged by the outlier program was implicit in the statistical methods used. With nine observations at the most, and with confidence intervals set around the 90 percent level, this is what would be expected if Student's t distribution is assumed.[18]

Another difficulty with this computer-oriented test was that in those cases in which the whole set of prices should have been suspect, rather than just one or two of them, the anomaly was usually missed completely. In particular, a very high dispersion about the mean might be a reason for which to be suspicious of the whole set of prices for a particular quality and country or, at least, of a substantial portion. Yet, in the outlier test such high variance would widen the interval within which acceptable observations could lie, and the problem might not be spotted.[19]

A further problem with the test was that its design was based on distribution functions that have been well studied, but that do not seem applicable when dealing with small price samples. Moreover, the distributions that seem to be appropriate vary according to category (e.g., different distributions would fit foodstuffs and services). In many cases the prices were distributed according to what could be considered a modal distribution. This was especially the case for categories affected by price controls.[20] In other cases, something approaching the normal curve would hold.

2.5.2 Objective outlier tests or subjective evaluations procedures? At the start of this data-handling stage, the criticisms leveled against *subjective* tests for detecting extreme values appeared quite convincing. Because of this, a determined effort was made toward the application of *objective* cleaning procedures that were unbiased in the selection of cutoff parameters. Yet, it has been seen that such outlier tests are subject to other troubles.

In sum, the computer-based outlier tests seemed to be useful for a preliminary checking of price dispersion but could not be considered definitive. Thus, the next step in the data-preparation process was the design of dispersion checks that would not be affected by the kinds of problems considered above.

[17] In some cases the outliers reflected the fact that the prices had been classified in the wrong quality slot (a price for rice, for example, being classified as superior rather than intermediate quality). On other occasions they were due to an inconsistent classification of brands or varieties (brand X toothpaste being considered sometimes as intermediate, and other times as inferior, quality). At other times, the unit of measurement was the source of trouble, with prices actually referring to units different from those specified. Finally, a number of mispunches were found, as well as prices that, even though screened out, appeared to be perfectly possible.

[18] In theory, one observation out of ten should have screened out in this case.

[19] If the observations are widely dispersed, but rather evenly distributed about the mean, the average price would remain largely unaffected. Thus, the irregularity may not be picked up by way of other checks.

[20] Under these conditions, an outlier might be perfectly valid. For example, all prices, except the ninth one, may be the same. The very low dispersion can easily frame such a price as an outlier.

It seemed that to comply with such requirements, the tests had to be based on a detailed examination of the computer listings.[21] The method finally adopted commenced with an examination of the price arrays for each quality, item, and country, as shown in the computer listings. A main objective of such scanning was to discern the dispersion patterns present in the data. As pointed out above, modal or quasi-modal distribution functions appeared to be prevalent, with a number of prices coinciding or closely grouped around a modal value. Once the appropriate central values were determined, ranges could be established around them, following some dispersion criterion.[22] The prices falling outside these limits would be considered suspect. The range was arbitrarily defined as *one-third* above or below the central value. This criterion provided a stiffer price dispersion test than the outlier program did.

The checking of computer listings, determination of central values, and the setting out of acceptable ranges were done with the aid of an assistant and a calculator. Once it was decided to go ahead with a detailed examination of the computer listings, the central-value ranges could be determined without much more work. For this reason, and also due to the flexibility required in the application of the various types of central-value ranges, the computer was not involved in the procedures.[23]

Thus, although conceivably the determination of the central value and its range could have been done by computer, the advantages that machines could bring were outweighed by flexibility requirements, and by cost considerations.[24] Another contributing factor was the circumstance that communication problems with the computer center were at their peak during the period in which these checks were required.

2.6 *The Editing of the Price Files*

Based on the tests described in the previous section, and on the checking of the sample price averages, a substantial number of prices were reclassified, adjusted, or deleted. In other cases, new information was requested from the field.

After these changes were incorporated into the card file, new computer listings were prepared and sent to the field, together with the unedited listings. The institutes were asked to appraise the editing done and encouraged to suggest other changes. As a result, new pricing was undertaken, either for testing purposes or to fill gaps in the survey.[25]

When the listings were returned to Brookings, it was found that in the great majority of cases the collaborating institutes agreed with the tentative editing done.

[21] These were being partially scanned anyhow, as part of the price-average checks.

[22] The mode was most frequently selected as the central value. However, in certain types of distributions, the median made more sense and was used instead.

[23] It was found that the great majority of the extreme values screened the second time around were also due to mispunches, incorrect quality classification, inconsistent relationships between brands and qualities, and the use of units of measurement different from those specified.

[24] The advantage of the machine in these routines lies in its greater accuracy and speed, together with the fact that it forces you to define explicitly the criteria used and to apply them consistently.

[25] These new prices referred to the original date of the inquiry: May 1968. Outlets were asked to provide prices corresponding to that month and year, although in certain cases current prices had to be deflated.

The other suggestions received from the institutes were generally incorporated into the data file. In some cases, though, their suggestions were quite foreign to the homogeneous editing procedures followed at Brookings and/or contradictory among the various countries involved. For the sake of overall consistency, no changes were introduced in these cases.

After the revision of the coding sheets and the critical examination of price averages and outliers, the first phase of the cleaning process was over. In it, the aid of the computer was quite significant, but the basic operations were mainly done by traditional methods. The computer speeded up the cleaning operations by quickly and accurately processing, calculating, and printing relevant information from the price survey (computer listings, averages, standard deviations, and so forth). Although it might appear that the computer should have been used even more, the special conditions of the project made this unadvisable. Some of these factors were: collaboration with heterogeneous research institutions resulting in diverse and/or poor reporting of the fieldwork; very broad country coverage; the infancy of the Brookings Computer Center; and the fact that much of the data was needed for only one kind of checking operation, which generally was not enough to justify keypunching expenses.[26]

2.7 *Analyzing Intercountry Price Dispersion*

The updated data were then put on magnetic tapes, from which price averages were again calculated and printed (see Table 3). A single geometric mean for all LAFTA countries was also calculated after converting the average prices into dollars.

TABLE 3
A SAMPLE LISTING OF COUNTRY PRICE AVERAGES FOR PARTICULAR ITEMS AND QUALITIES, 1968 ECIEL BENCHMARK
[in dollars][1]

Product	Product Code	Country Code	Superior Quality	Intermediate Quality	Inferior Quality	All Qualities
Beef, first class	1111013	01	–	0.84	0.77	0.81
" "	1111013	02	–	0.91	0.78	0.84
" "	1111013	03	–	1.00	0.62	0.79
" "	1111013	04	–	2.05	1.75	1.89
" "	1111013	05	–	1.03	0.86	0.94

[1] The official exchange rates used for conversion purposes are taken from: International Monetary Fund, *International Financial Statistics*, Washington, D.C., 1969.

Two basic types of dollar averages were calculated: one, using the official exchange rates for conversion purposes; and the other, utilizing the ECLA purchasing-power parity rates corresponding to the private consumption sector.[27]

[26] An example of the computer-center problems that had to be faced was the way in which the editing and updating of the price file was conducted in the initial stages. These had to be performed on the cards, because of delays in writing an update program. This, in turn, resulted from the needlessly complex tape structure proposed by computer specialists for the data file.

[27] On the latter rates see U.N. Economic Commission for Latin America, *op. cit.* The ECLA rates were extrapolated to May 1968.

With this material, matrices of price relatives by quality and item were formed. The price relatives converted by way of the ECLA rates were used in checking intercountry price dispersion. These rates were utilized because they equalize the purchasing power of money over consumer goods in the countries involved, and therefore their effect on the price relatives for individual products could be considered to be neutral. In contrast, official or other kinds of exchange rates distort the dollar conversions and, thus, the price relatives.[28]

An example of such a price relative matrix for a few countries is given in Table 4. In that table, each row is the set of price relatives with the country in the stub as a base. The last row is headed by the overall LAFTA average and provides the price relatives of the various countries with respect to it. Each column presents the price relatives of the particular country heading it with respect to all possible bases.

TABLE 4
A Sample of the Matrix of Price Relatives for Beef, First Class, from the 1968 ECIEL Benchmark for LAFTA Countries

	Numerator				
Denominator	Argentina	Bolivia	Brazil	Chile	Colombia
Argentina	1.00	1.04	1.05	2.53	1.54
Bolivia	0.96	1.00	1.01	2.42	1.48
Brazil	0.95	0.99	1.00	2.40	1.46
Chile	0.39	0.41	0.42	1.00	0.61
Colombia	0.65	0.68	0.68	1.64	1.00
LAFTA average	0.62	0.64	0.65	1.55	0.95

These computer-generated matrices, as well as the country price averages referred to earlier, were used to check intercountry price dispersion. The basic assumption behind such tests was that even though technologies, factor costs, demand conditions, economic policies, and so on, differ from country to country, the prices of particular products could not be very far apart.

A test was designed which operated on the price relatives, spotting wide deviations from the LAFTA average. Extreme price deviants were defined as those that were over twice, or below one-half of, the LAFTA average. However, further confirmation of the extremeness of these deviants was sought by looking at other relevant parts of the price-relative matrices. Moreover, a quick reference to the price averages provided additional indications about the plausibility of these outlying prices. In the end, many of the extreme prices were left in, some because they appeared justifiable, and others because their importance in total consumer spending was found to be rather low.

Although the first part of these checks could have been done by computer, the fact that the whole matrix of price relatives would probably have to be consulted later spoke against any such use of the machine. The computer was not found to be economically efficient in providing a quick scanning of a varying subset of the price-relative matrix.

[28] On this point, see Joseph Grunwald and Jorge Salazar-Carrillo, *op. cit.*, pp. 242–248.

The main justification for scanning the matrix of price relatives is that it is difficult to determine by rigid rules which prices are actually deviants. Just one extreme observation can give rise to an irregular pattern of price relatives. It is necessary to look at various sections of the matrix, rather than just the row headed by the LAFTA average, in order to have a clear picture of the problem.[29] Unfortunately, the parts of the matrix which are relevant vary with the product in question.

Apart from these efficiency considerations, it was felt that if the alternative of writing an algorithmic program had been chosen, it would have taken longer. Furthermore, a sizable amount of effort and resources would be spent in communicating the operations involved to a programmer, and in checking his program and output.[30] On the other hand, one important advantage of computerizing such operations would be that it forces the spelling out of an explicit and consistent procedure for determining which prices should actually be mistrusted. However, the disadvantages outweighed the advantages in this case, and the whole series of tests were done by hand.

A list of prices to be verified in the various countries emerged from these tests. For such purposes, computer listings of the edited price files, together with information about the outlying price relatives, were sent to the institutes. About three-quarters of the suspected price deviants were changed in the verification process that ensued, with all price modifications referring back to May 1968. At the end, the price file was again updated to reflect these changes.

In all, the net contribution of computer operations appears to have been somewhat larger in this second phase of data managing, with the computations it performed and printed being central to the testing done. This seems to reflect the fact that once a reasonably clean price file has been created, the contribution of the computer to succeeding data-handling tasks is bound to increase. Two other determining factors were: (a) increased reliability in the operation of the computer center; (b) the experience gained by the research staff in the use of the computer.

2.8 *Cleaning the Other Variables*

Information on a number of qualitative variables was also collected in the ECIEL benchmark survey for 1968. These characteristics pertain to each of the prices obtained. The principal ones are:

(a) name of the store;
(b) type of outlet (e.g. supermarket, small grocery store);
(c) income stratum of its clientele;
(d) brand and model or variety of the item priced;
(e) quality classification; and
(f) unit of measurement.

[29] To clarify further, one or a few high price-relatives may make some others appear low. Determining where the problem lies in cases like these may require looking at several rows and columns of the matrix or even to fall back on the country price averages. It is interesting to note that only about 4 percent of the total amount of country price averages were considered questionable at this stage.

[30] An assistant working in the project would be expected to grasp the essence of the test in a relatively short period of time.

Evaluation of the quality classification given in the field has been dealt with in previous sections. Apart from the quality indicators, only two other price attributes were processed and cleaned to some extent: the type of outlet and the income stratum of its clientele. These data were kept in the price file referred to above.[31] The burden of verifying such data was basically placed on the institutes. The central coordinating unit at Brookings attempted to obtain missing information from the institutes and insisted on the verification of the data received. Computer listings were sent to the institutes at various times, partly with this objective in mind. Any editing done by the institutes, or any new information sent, was incorporated in the price file.

The response to such editing tasks was mixed. Some countries thoroughly cleaned the information on type of store and income stratum, while others did so partially. The main categories, like foodstuffs, clothing, and rent are quite complete. Others, like consumers durables and services, present some problems. In all, most of the data can be utilized if care is exercised.

A similar type of checking was carried out with the other information mentioned, even though these were not included in the price file. This was done principally because such data were useful in the process of cleaning the price information. Because of these needs, the cleaning of these other attributes was relatively successful. However, the information on brand or variety, unit of measurement, and name of outlet did not appear useful for other research purposes and it was not subjected to systematic editing.

2.9 *The Calculation of Price Indexes*

Once the data were clean, the stage was set for the calculation of index numbers and their evaluation.[32] Of the different index formulations considered, the Walsh index with geometric weights seemed to be the most appropriate.[33] This index is a geometric price-relative formulation which geometrically crosses the expenditure weights of the countries involved.

The consumer expenditure weights required will come from a parallel ECIEL study, which is also coordinated by Brookings. This study is based on consumer expenditure surveys in the most important urban areas in LAFTA.[34] Because the weights generated in that project were ready for only a few countries, an alternative set was required in order to proceed with the construction of preliminary price indexes. The information utilized by the Economic Commission for Latin America in their 1960–1962 study was used for this purpose.[35]

[31] Only a limited number of store types and three income strata were specified in the study. These were kept homogeneous across countries.

[32] Of course, the items and qualities to be used in the calculations had to be determined first. As mentioned above, 416 items were included in the price survey, with three qualities priced for each. About 30 percent of these 1,248 qualities had to be excluded from the computations, because of omissions, errors, unreliability, and other data problems. However, in terms of the items themselves the situation was much better, with only 3 to 4 percent omitted from the calculations, which is a pretty low rejection rate.

[33] The advice of Richard Ruggles was most helpful in reaching this decision. For a comprehensive treatment of this topic, see Richard Ruggles, "Price Indexes and International Price Comparisons," in William Fellner et al., *Ten Studies in the Tradition of Irving Fisher*, New Haven, 1967.

[34] See the paper by Howe and Villaveces that is included in this volume.

[35] See U.N., Economic Commission for Latin America, *op. cit.*

The computer was relied upon entirely in the calculation of the price indexes. It was asked to relate the matrix of price relatives with the expenditure weight scalars according to the Walsh index-number formulation. This formula could be written simply as:

$$\prod_{i=1}^{n}\left(\frac{P_{ij}}{P_{ik}}\right)^{W_{it}},$$

where P indicates prices, i are products and j and k countries, and W_{it} is the geometric average of the country weights for every product. Place-to-place indexes were calculated for overall private consumption and for more than a hundred of its major categories and subcategories.

Apart from their overall usefulness for research, these indexes and subindexes can be used for further checks on the data. This is especially the case with results at lower levels of disaggregation. In fact, some of the preliminary subindexes suggested that there were problems in some categories (e.g. tobacco products, fuel and utilities), with dispersion being much higher than expected. These were scrutinized by researchers from the collaborating institutes in one of the semiannual ECIEL seminars, and some rechecking was done in the field when necessary. This required a minor updating of the price file, which was done by computer at this point.

The results of these tests were generally positive. The index numbers and the country rankings generated seemed quite reasonable. The overall index and subindexes did not diverge widely from the ECLA results in 1960, and the country rankings were quite similar.[36] As these surveys were largely independent, it can be concluded that these studies tend to support each other, and that a high degree of confidence can be placed in their findings.

Aided by the computer, several recalculations of the Walsh price index have been undertaken after the checking was completed. The indexes for overall private consumption and its main categories have been presented and analyzed in two recent papers.[37] The indexes corresponding to the various subcategories have not yet been published.

Table 5 shows a price-index matrix for a particular subcategory (preserves) as an illustration of the format in which these results are available. Some of the countries, as well as the LAFTA average, are excluded from the sample matrix shown. The structure of this table is similar to that of the price-relative matrix presented above. Each row is headed by the base country, while each column is headed by the other country.

The same process has been adapted to comparisons between pairs of countries within LAFTA. However, the index formulations used are different. The traditional Laspeyres and Paasche index numbers were calculated in this case, together with their geometric mean: Fisher's Ideal Index. As an illustration, some results from the Venezuela-Colombia binary comparison are included here. It should be noted

[36] On these points see Joseph Grunwald and Jorge Salazar-Carrillo, *op. cit.*, pp. 265–273.

[37] See Joseph Grunwald and Jorge Salazar-Carrillo, *op. cit.*, pp. 265–275, and Jorge Salazar-Carrillo, "Price, Purchasing Power and Real Product Comparisons in Latin America," pp. 128–131. The results presented in these papers differ. The latter incorporates the latest updates of the price file. Also, the weights used in the first paper refer to 1960, while estimates for 1968 were utilized in the second.

TABLE 5
MATRIX OF INTERNATIONAL PRICE SUBINDEXES FOR PRESERVES FOR SOME LAFTA COUNTRIES. IN 1968[1]

	Numerator				
Denominator	Argentina	Bolivia	Brazil	Chile	Colombia
Argentina	1.00	1.76	1.38	1.68	1.32
Bolivia	0.57	1.00	0.79	0.95	0.75
Brazil	0.72	1.27	1.00	1.21	0.95
Chile	0.60	1.05	0.83	1.00	0.79
Colombia	0.76	1.33	1.05	1.27	1.00

[1] The official exchange rates used for conversion purposes are taken from: International Monetary Fund, *International Financial Statistics*, Washington, D.C., 1969.

that these results, in contrast to those referred to previously, are final rather than experimental. This is because for these countries the weights used were taken from the ECIEL consumer-expenditure surveys.

In Table 6, both the Laspeyres and Paasche indexes for Colombia, with Venezuela as a base, are presented. This is done for overall private consumption and its major categories. As can be seen, the discrepancy between the Laspeyres and the Paasche is quite wide, with the Laspeyres being 63 percent above the Paasche for overall private consumption.[38] The Ideal Index of Fisher, a relatively unambiguous estimate of the price difference between these countries, is also shown.

TABLE 6
BINARY PRICE COMPARISONS BETWEEN COLOMBIA AND VENEZUELA FOR OVERALL PRIVATE CONSUMPTION AND ITS MAJOR CATEGORIES, WITH VENEZUELA AS THE BASE (1968)[1]

Categories of Goods and Services	Laspeyres	Paasche	Fisher
Food	0.87	.66	.76
Beverages	0.59	.52	.56
Tobacco	0.42	.33	.37
Clothing	1.19	.70	.91
Rent	0.52	.43	.48
Utilities	0.35	.30	.32
Durables	1.20	.76	.96
Nondurables	1.01	.68	.83
Services	0.87	.36	.56
All goods and services	0.83	.51	.65

[1] The official exchange rates used for conversion purposes are taken from: International Monetary Fund, *International Financial Statistics*, Washington, D.C., 1969.

As can be seen, consumer prices in Colombia are about two-thirds those of Venezuela at the official rate of exchange. This varies by category, but Colombia's consumer prices are lower throughout if the Fisher formulation is used. This result is not surprising. In terms of private consumption, Colombia has relatively low

[38] This is ultimately determined by the inverse correlation generally existing between prices and quantities.

prices if compared with other Latin American nations, while Venezuela has the highest prices of the group.

It should be stressed that the computer was all important in this last data-handling stage. Not only did it allow for quick and accurate calculations, as well as numerous disaggregations, but it facilitated the computation and testing of alternative formulas. In terms of cost 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.[39]

Thus, the utilization of computers in this last step of data preparation was very heavy. All calculations were done with the help of computers. It would seem that it is in the calculation stage that computer operations enjoy their greatest comparative advantage.

Section 3

3. *Computer Use in the ECIEL Price Study and the Usefulness of Computers: An Overview of Benefits and Drawbacks*

The 1968 ECIEL benchmark involved large masses of data. These had to be processed, organized, checked, cleaned, and used for the computation of the varying number of estimates required in the study. The computer was the backbone of these data-handling tasks, with some of the operations involved being entirely computerized.

A definite pattern can be noted in the ECIEL study computer usage. At the beginning, when the data were being processed for storage and organized as a file, the computer was heavily relied upon. The degree of computerization was relatively less intense in the cleaning and editing stages, even though it picked up somewhat at the end. During the final measurement and estimation stages, virtually all of the work was done by machine. The pattern seems to be ultimately dependent on the characteristics of the data at each stage and/or the type of work to be done on them.

The experience gained in handling the data gathered in the ECIEL benchmark price collection has shown that some of the supposed advantages of computers (see section 1) do not necessarily materialize. The advantage of increased quickness and accuracy was to a large degree offset by the time researchers had to spend in verifying that the instructions had been well transmitted throughout the chain of command (senior researcher, junior researcher, programmer, computer). A very careful and time-consuming check of the output was also necessary. This can be a formidable task, given the tremendous masses of results that the computer generates in projects of this nature.[40]

The avowed quickness of the computer was found to depend on the operation being executed. The computer can process and organize large bodies of data in little time. Nevertheless, in other tasks it does not perform as efficiently. Moreover, in every task, the computer is limited by the fact that some of its components

[39] Communication problems still remained. Computer programs and the output they generated had to be screened extra carefully.

[40] On this point refer to the information overload concept developed in the Howe-Villaveces paper included in this volume.

may act as bottlenecks. In the present study, things like keypunching, card reading, programming, and so on, on occasion created major delays that prevented the efficient use of the machine, diverting some work to the traditional method of hand calculation.

One of the more substantial contributions of the computer to the project lay in the idea of a data bank. If computers were not available, the possibility of the ECIEL data being used for other research ventures, or by other researchers, would be much smaller. Computer operations make the researcher aware of the potentialities for further use of the information and enriches his overall research product. At the same time, the requirements of data banks place certain demands on data organization, forcing the development of an efficient file structure aimed at rapid data retrieval.

The major contribution of the computer to the experience described here was in performing the final set of calculations. The research phase of the study has not been completed yet, but it can be expected that because of its computational advantages, the computer will handle nearly all of the remaining operations. It is also foreseen that the computer will permit the use of the price data file for varied experimentation and hypothesis testing, well after the main results have been presented.

On the other hand, it must be recognized that the use of the computer was, in some respects, burdensome. From the start, a project deeply plagued by lack of manpower had to devote a substantial amount of resources to mastering the essentials of computer operations. Discussions and relations with programmers and other computer personnel seem to have taken a disproportionate amount of time. One important lesson stemming from the study is that for sufficiently large projects, it is preferable to hire programmers directly, rather than to rely on those employed at the affiliated computer center. This is to a large degree determined by the lack of service orientation on the part of many computer centers, which seem to have a tendency to become absorbed with systems problems, shortchanging the users in the process.

Other problems with computer use involve the partial loss of independence and difficulties in communications. The latter are more important in day-to-day operations, the former in the long run. Communication problems with computer personnel can be quite wasteful.[41] The experience of the study described was rather negative in this respect. Complicated theoretical systems were designed never to be put into use, while in other cases, computer-center programmers spent months with the project without producing results. This not only slowed down progress, but also affected computer usage by diverting tasks away from it.

Loss of independence seems to be potentially dangerous for projects with large bodies of data. Once a file has been registered at a computer center, and a set of data-management programs has been developed, a researcher faced with unsatisfactory computer operations has few options available. The data- and

[41] One interesting point that seems to hold frequently is that the more sophisticated the programmer, the more complicated his solution to simple problems. The first system designed to handle the ECIEL price data would have required $200.000 for its full development. Due to the author's insistence, a less complex system was made operational by a lower-level programmer, effecting considerable savings.

program-conversion problems implied, the costs of establishing new computer relationships, the difficulties of rationally considering the alternatives (buying commercial time or exploring other possible computer linkups) tend to bind him to the status quo.

Section 4

4. *The Use of the Computer in Price Research: Final Conclusions*

After considering and weighing both the beneficial and harmful consequences of the computer for the ECIEL price benchmark, it appears that the overall impact has been positive. However, the strict controls and restrictions placed on computer usage had much to do with this outcome.[42] An additional contributing factor was the realization that programs, computations, and results had to be meticulously checked. Finally, another determining element was the continuous effort made to improve the understanding of computer operations and to better communications with computer-center staff.

The main conclusion of this paper is that the use of computers in price benchmarks in particular, and in price research in general, would be quite advantageous and thus should be expanded. However, their net contribution is likely to be much larger if the qualifications discussed above are taken into consideration in planning the role of the computer in each project.

Another important point arising from the experience gathered in the ECIEL study is the observation that the use of the computer in each task is substantially affected by factors that go beyond the computer per se. The quality of the computer center being dealt with, and of the programmers helping out, are just two of these factors. Two additional elements are the researcher's experience in computer work and his ability to relate to computer specialists. Thus, if each task is considered as output, the input coefficients corresponding to computer services would probably vary substantially, depending on these factors. Thus, things like having programmers attached directly to the projects (rather than using those on the computer-center staff) and keeping open the possibility of switching to alternative computation facilities are well worth pondering by researchers in the area of prices.

There is reason to be confident that computer science and management will improve fundamentally in the years to come. In particular, the development and canning of cleaning and editing programs amenable to price research should be expected. An improvement in the management of computer aspects of research projects, together with greater emphasis on services and user needs, also seems likely.

Moreover, researchers can help achieve a more efficient utilization of the computer setup, particularly if they are working in the same field. This could be done by an agreement to specialize in different aspects of computer work and to share the results with one another, as well as by exchanging programs and other

[42] As has been seen, there was no blanket endorsement of computer usage in the project. Its advisability in each task was carefully examined, with the knowledge that the net advantage or disadvantage of computers over traditional methods varies greatly from one application to another.

computer information, so as to avoid the enormous duplication of effort presently prevailing. To implement this proposal, it would be necessary to ask programmers and computer centers to tailor their work so as to facilitate its transfer to other computer centers.[43]

Hopefully, the winds of change will blow in such directions, because deriving reliable, up-to-date results for studies of the nature described above can be accomplished only with the computer. Being able to use its capabilities more efficiently may open new research frontiers in the area of prices, in general, and for price benchmarks, in particular.

[43] A start could be made by encouraging linkages among computer centers on a selective basis.