

MARYLAND INTERINDUSTRY FORECASTING PROJECT

Research Memorandum No. 23

October 14, 1969

**ADJUSTMENT AND AGGREGATION OF HISTORICAL SERIES FOR
MANUFACTURER'S SHIPMENTS AND COST OF MATERIALS
1947-1966**

by

Theophilos Terrovitis

As the MIFP converts to the detail offered by the 1963 input-output table, we need historical series in constant dollars for output and cost of materials of the new sectors. This memorandum reports how such series were derived for the manufacturing industries.

Since the four-digit industries are the "molecules" of government statistics, we begin by constructing the output series for them. Then we describe the I-O sectoring plan in terms of four-digit industries and, on the computer, aggregate the four-digit series to the I-O sectors. The same industry description and aggregation program can be used in preparing the import and export statistics, and for adding each new year's data to the historical series.

Our constant-dollar output series are derived from the work of the Real Product section, National Economics Division of the Office of Business Economics. Mr. Jack Gottsegen of OBE has been most generous in helping us to understand and use these data. They provide the value of shipments and the cost of materials, both in current and constant dollars. The data fall into three periods. For 1947-1957, they are based on the 1954 SIC and were obtained from work sheets. The second-period data, for 1958-1962, are based on the 1958 SIC and are available on tape. The final period, 1963-1966, and the 1967 deflators, was covered by a tape on the 1963 SIC basis. To obtain our output series, therefore, we had first to adjust all the series to the same SIC base. This adjustment was done at the four-

digit level. Then the four-digit series were aggregated. This memorandum describes the methods and programs used for this adjustment and aggregation for the output series and the cost-of-materials series.

I. OUTPUT SERIES

A. Adjustment to 1963 SIC Classification

This adjustment required two steps:

- 1) The data for 1947 to 1957, which are given in the 1954 classification, were adjusted to the 1958 classification.
- 2) The adjusted data for 1947-57 and the 1958-62 data, which are given in the 1958 classification, were adjusted to the 1963 classification. The entire series, 1947-66 are then on a common basis.

For the first adjustment, the available information was the Distribution of the Value Added in 1958, given in Appendix C of the 1958 Census of Manufactures. This appendix gives the part of 1958 value added, VAP_{ij} , of each old SIC industry i going into each new SIC industry j . It also gives the total value added in the old industry i , VAO_i , and the total in the new industry j , VAN_j . Let X_{it} be the shipments by old industry i in year t , $t = 1947, \dots, 1957$. Then we took as the shipments of new industry j , A_{jt} , the expression

$$A_{jt} = \sum_i (VAP_{ij} / VAO_i) * X_{it} * \frac{VAN_j}{\sum_i VAP_{ij}}$$

The last term, $\frac{VAN_j}{\sum_j VAP_{ij}}$, should be 1.0 if all parts of the new SIC are accounted for. In some cases they are not, so it effects a slight upward adjustment.

All of this adjustment has been programmed and executed on the computer.*

*The adjustment program uses two decks of cards and works as follows:

- a) The computer reads first the time series (1947-57) deck of the Value of Shipments for the old four-digit SIC industries and the corresponding Value Added (VAO) for 1958. The SIC numbers must be in ascending order.
- b) Then it reads the first card from the second deck of cards, which contains:
 - 1) The new four-digit SIC number, NEWSIC, and its Value Added (VAN), and the number of old sectors in the new one, NC.
 - 2) The old SIC numbers which contribute to the new one, ISIC(I) and how much each sector contributes to the new SIC number.

When the computer reads this first card, it searches in the first deck of cards, to find the first of the old SIC numbers of this card. In the search, the computer starts from the first number and checks every twentieth number until it has passed the number for which it is looking. Then it backs up one hop and checks every number until it finds the one it wants. When it finds it, it takes the ratio VAP_{ij}/VAO_i and multiplies it times the series of the value of shipments^{ij} of the old SIC number, and adds the results into X_j , the value of

(continued)

This same procedure has been used to adjust the (1947-62) time series to conform to the 1963 version of the SIC. The only difference is that the relationship of the Value of Shipments of the old to the new SIC numbers (1963) was used, instead of the Value Added relationship.

B. Aggregation

The adjustments of the data described above were intermediate steps to obtain comparable series for each four-digit SIC industry. The final step is to aggregate these four-digit industries to the sectors used by the 1958 and 1963 input-output tables.

This aggregation was done on the computer by a program that was written for this problem but can be used for aggregating exports, imports, or employment as well. It is described in the Appendix of this memorandum.

II. COST OF MATERIALS

The cost of materials data were obtained from the same sources from which we obtained the output series, namely, from the Department of Commerce, Office of Business Economics. Naturally, we started from the four-digit SIC numbers. For 1947-57 they were obtained from work:

Footnote continued

shipments of the new industry. Then it takes the second old SIC number on the first card and repeats the same procedure. So it continues until it finishes all NC of the SIC numbers on this card. Finally, it multiplies the X series by the VAN/EVAP ratio, and the new SIC number is ready. Then the computer reads the second, third, etc., card from the second deck and repeats the same procedure until all are done.

sheets and are based on the 1954 SIC. For the next nine years, 1958-66, they were taken as the difference between value of shipments and value added in the Census and Annual Surveys of Manufacturers. They are based on the 1958 SIC, through 1962 and then on the 1963 SIC. To obtain the four-digit SIC on a common basis, we adjusted all the series to the 1963 basis by a two-step adjustment, as in the output series. Finally, the cost of materials by I-O sectors were obtained by using the aggregation program.

One difference from the output-series-computations, lay in the way we deflated the cost of materials to 1958 dollars for the period 1958-66. For 1947-57 we had from OBE, the four-digit cost of materials in 1958 dollars. For the next nine years, 1958-66, we made deflators for the cost of materials in each industry by taking a weighted average of the output deflators for the supplying industries. For weights, we used the coefficients of the corresponding column of the input-output A matrix (adjust the sum to 1.0). Finally, we divided the current-dollar values of the cost of materials in the I-O sector by the deflators so obtained. In this way, we estimated cost of materials in 1958 dollars.

* * * *

Punched-card decks of Shipments in current and constant dollars have been prepared for four-digit SIC sectors, manufacturing sectors of the 93-order I-O table and of the new 170-order table. Cost-of-materials in current dollars are available for all these sectoring and, in constant dollars, for the 93-order sectoring.

APPENDIX

The Aggregation Program

Written in cooperation with Bertram Wolfe

The aggregation program transforms a vector classified by SIC code into one classified by input-output sector and attaches the results as the latest observations in the historical time series in the forecasting program. It can aggregate output, cost of materials, exports, imports, employment, or what-have-you with complete indifference.

The aggregation can be divided conceptually into three parts. The first section of the program reads into memory all the SIC numbers and corresponding series which will be needed to construct the I-O series.

The second part of the program reads the identification number of the first I-O sector for which we want a series, the count of how many SIC categories will be used to make up this I-O sector, and then the specific SIC identification numbers of these categories.

The third part of the program actually performs the calculations. In general terms, this process searches for the SIC outputs which go into the I-O sector and then adds or subtracts the SIC values to get the series for the I-O category.

In order to run this program, the user needs to understand only the first and second sections, which deal with the input of data, and the form in which the output will appear. Nevertheless, the details of the calculations will be discussed briefly after we explain the specific operation instructions.

The data deck for the program should be set up as follows for running the Univac 1108.

1. Title Card. Anything Punched in less than 78 columns is printed.
2. Control card with NCODE, ISTOP, LIRIO, NPGO, NPSTART, and NPSTOP.
3. Data deck ONE which is the previously prepared I-0 series followed by a 99999999999999999999999999999999 card.
4. Data deck TWO which is the SIC series followed by a 99999999999999999999999999999999 card.
5. Data deck THREE which lists the SIC sectors in each I-0 sector.
6. Fin card.

The first card is a title card. The first 78 columns will be printed on the output. The second card is punched with the control constants:

NCODE, ISTART, ISTOP, LIRIO, NPSTART, NPSTOP. These constants are:

NCODE: an identification code which will be punched on the output cards and which tell what kind of series it is.

ISTART: the last two digits of the starting year for a previously prepared I-0 series.

ISTOP: the last two digits of the last year of a previously prepared I-0 series.

LIRIO: a 1 if there is a previously prepared I-0 series, otherwise, 0.

NPGO: the last two digits of the first year of the new data.

NPSTOP: the last two digits of the last year of the new data.

If LIRIO is 1, data deck ONE, the previously prepared I-0 series must be present. The format used for reading this deck is (I2,I2,5X,10F7.1). The I-0 sector number is punched into the first integer field, NCODE is punched into the second I field and the value of the series in successive periods is punched into the F fields which follow.

Following deck ONE, or if LIRIO is zero, following the control card, comes data deck TWO, read with the format (10X,I5,10F6.0/(15X,10F6.0)). The second data deck gives SIC numbers and their corresponding time series. An entire twenty year series for one SIC can be punched into two successive cards. For longer series, the data would have to be punched into three successive cards. The SIC number is read from the first card of each series. It is written as a five digit number, e.g. SIC 23 appears as 23000 .

At the end of the SIC series deck, it is necessary to insert a card which has 9's punched in columns one through fifteen. It is also important to keep the SIC series deck in order by SIC number; otherwise the search routine will not work.

Deck THREE tells which SIC series should be added together to get the I-O series. The format for the third deck is (2I5,10I7/(10X,10I7)). The I-O sector identification number is punched into the first five columns (right justified). In columns 6 through 10, the number of SIC numbers used to define the I-O sector must be punched. The rest of the card contains this SIC description of the sector. If the SIC number is positive, the series for that number will be added into the I-O series in each time period. If the SIC number is negative, the SIC series will be subtracted from the I-O series. In this way, the program makes provision for cases where the I-O sector is most easily described as, say, a two-digit SIC less one or more three-or-four-digit-SIC numbers.

For example, if I-0 sector 26 is equal to SIC 225, 23 except 239, and 3992, its description card would appear as follows:

26 4 22500 23000 -23900 39920

When the data decks are properly organized and a copy of the source program and system control cards are added, the desired calculations can be carried out by the computer. After decks ONE and TWO have been read, the computer reads the description of an I-0 sector, initializes the value of the series for that sector to zero, and begins the search for the first SIC number of an I-0 sector. It compares every twentieth SIC number in deck 1 with the SIC number it is hunting for. When it finds that it has passed the number it wants, it backs up twenty numbers and looks at each one. After finding the series and adding it to (or subtracting it out of) the I-0 series it is building up, it starts forward from there to look for the next SIC in that I-0 sector. It is therefore very important that the SIC numbers be in order in deck ONE and in each I-0 sector description. After the machine has found the last SIC sector in the I-0 sector, and done the necessary additions or subtractions, it prints and punches the series for that sector. Then it reads the next I-0 description card and repeats the process. The repetition continues until the machine runs out of data cards.