INTERINDUSTRY FORECASTING PROJECT

UNIVERSITY OF MARYLAND

Research Memorandum No. 2

HOW TO PURIFY AN INPUT-OUTPUT MATRIX

By

Clopper Almon

5 July 1967

HOW TO PURIFY AN INPUT-OUTPUT MATRIX

I. THE SECONDARY PRODUCT PROBLEM

One of the perennial trouble makers in input-output is the secondary product. Examples of secondary products include stampings cut in a plant primarily engaged in rolling aluminum, breakfast cereals made in a plant whose primary product is paperboard, or lumber sawn in a plant whose main product is furniture. In the 1958 matrix, such products were treated by a transfer. For example, the table shows a $19 million "sale" of furniture to lumber, which is really a transfer of lumber made in furniture plants back to its home industry. The number of such secondary entries will increase substantially as the number of sectors increases. These entries are not only confusing when looking at the table, but they can also distort the outcome of its use. If we asked what would be the impact of an increase in residential construction, the big lumber user, this furniture-to-lumber secondary entry would lead to an immediate increase in the output of the furniture industry and thereby to higher demand for textiles, wire, paint and varnish, and glue. Moreover, because the establishments in the Structural Metals product sector make some furniture, which is transferred to the Furniture sector, the demand for lumber also generates demand for structural steel to "go into" the lumber. Such distortions are probably not large, but they are disconcerting and have led to a desire to "purify" the matrix. This paper describes one purification rite.

II. THE PURE ACTIVITY MATRIX

The simplest approach postulates that each product is made by the same process, no matter what kind of establishment makes it. The vector of inputs into any establishment is therefore just a linear combination of the pure processes for the items it makes. In symbols

(1) $$f_i = Mp_i \qquad i = 1, \ldots, n.$$

where:

f_i, a column vector, is the transpose of the i^{th} row of the primary flow matrix, that is, f_{ij} equals the purchases of product i by establishments in industry j.

M is the product mix matrix: m_{ij} = the fraction of product j made in establishments in industry i. (The columns of M sum to 1.0).

p_i is the transpose of the i^{th} row of the pure flow matrix; p_{ij} = inputs of product i into product j.

Equation (1) can be solved for p_i by

(2) $$p_i = M^{-1}f_i \qquad i = 1, \ldots, n.$$

The p_i vectors found by this formula will generally have some small negative elements and some small positive entries where the corresponding f_i vectors have zeros. The former are clearly nonsense; the latter are at best dubious and certainly inconvenient, especially as we move toward larger matrices for which we will wish to store only non-zero entries in our computers.

By analysing the solution of (2), we shall find a way to put a stop to this nonsense.

III. THE ITERATIVE INTERPRETATION

In the rest of this paper, we shall drop the subscript i from f and p and understand that the f and p in an equation all have the same i subscript. Equation (1) may be then written

$$0 = -M_p + f \ ,$$

and by adding p to both sides we obtain

$$p = (I - M)_p + f \ . \tag{3}$$

The column sums of the absolute values of I - M will be less than 1.0, save in the unlikely case in which less than half of a product is produced by establishments primarily engaged in its production. The iterative process for the solution of (3) will therefore converge. In this process, we take as a first approximation of p,

$$p^{(0)} = f$$

and then define successive approximations by

$$p^{(k+1)} = (I - M)\ p^{(k)} + f \ . \tag{4}$$

To see the economic interpretation of (4), let us write out the equation for the use of a product, say electricity, in making another product, say j.

$$p_j^{(k+1)} = f_j - \sum_{\substack{\ell=1 \\ \ell \neq j}}^{n} m_{j\ell}\, p_\ell^{(k)} + (1-m_{jj})\ p_j^{(k)} \tag{5}$$

The first term on the right of (5) tells us to begin with the electricity purchases by the establishments in industry j.

The second term directs us to remove the amounts of electricity needed for making the secondary products of those establishments, using our present estimate $(p^{(k)})$ of the technology of those products. Finally, the last term causes us to add back the electricity used in making product j in other industries. The amount of electricity added by the third term is exactly equal to the amount stolen, via second terms, from other industries on account of their production of product j:

$$(1-m_{jj})\ p_j = \sum_{\substack{\ell=1 \\ \ell \neq j}}^{n} m_{\ell j}\ p_j$$

since $\sum_{\ell=1}^{n} m_{\ell j} = 1$.

IV. PUTTING A STOP TO THE NONSENSE

It is now clear how to keep the negative elements out of p. When the "removal" term, the second on the right of (5), is larger than the "primary use" term, the f_j, we simply scale down all components of the removal term to leave a zero balance. Then instead of adding back the "total-stolen-from-other-industries" term, $(1-m_{jj})p_j$ all at once, we added it back bit-by-bit as it is captured. If a plundered industry runs out of electricity with only a third of the total amount of plundering claims satisfied, we simply add only a third of each plundering product's claim into that product's cell in p.

V. IMPLEMENTATION FOR 1958

This procedure has been programmed and applied to the expanded 1958 matrix. The result is available as IFP Listing 2.

This listing shows the programs used for the calculation, and a row-by-row listing of the primary matrix, the purified matrix, and the secondary matrix. The portion of the program which performs the purification is SUBROUTINE PURIFY. Comment statements within this subroutine should make it self explanatory. In particular, it is noteworthy that it requires only non-zero elements of M to be stored. The 1958 matrix had less than 2000 such elements. This subroutine is therefore usable on the 7094 for matrices with up to five times more non-zero elements than the 1958 matrix had.

The primary matrix used in these calculations differs from the primary matrix obtained from the OBE tape (See IFP Research Memo 1) in three ways:

(1) It has the added detail for the food, non-ferrous metals, and utilities which was published in the Survey of Current Business, April 1966. Secondary flows for this further detail were supplied by OBE.

(2) OBE calls all the inputs into the two dummy industries, Business Travel and Office Supplies, secondary transfers. We have treated them as primary sales.

(3) Scrap and By-Products are put into a separate row (number 94) with no corresponding column. Positive entries in this row show net use of scrap or by-products; negative entries show net production. Naturally, the non-negative constraint of the purifying program was not applied to this row.

Comparison of the purified matrix with the primary matrix shows that in the vast majority of cells little change was made. A few zeros appear in the purified where the primary had entries. Most of these make good sense. For example, the primary matrix shows a small purchase by Aircraft from Ships, Trains, Trailers, and Cycles; the purified shows a zero. The purification process found that this purchase was more than accounted for by the purchase of parts needed for the "Ships, Trains, Trailers, and Cycles" made in the Aircraft establishments. Other zeros reveal inaccuracies in the original table. For example, we find a zero for the sale of electricity to Printing and Publishing and to TV and Radio Broadcasting, both manifest absurdities. The trouble is that the purified Business Services electricity use is $250 million for an output of $23,269 million, of which 22 percent was advertising transferred in from Printing and Publishing and 6.5% was transferred from TV and Radio. 22 percent of $250 millon is $54 million, $2 million more than Printing and Publishing's primary purchase of electricity. How did Business Services electricity use get up to $250 million? The first $186 million was a primary assignment; from that base it was easy to grab all of the electricity input into Printing and TV and Radio to bring the total to $240; the remaining $10 were collected from other small secondary producers of Business Services. The source of the difficulty is clearly the $186 primary assignment of electricity to Business Services. Upon inquiry, we learned that OBE made the electricity allocations in the service sectors in proportion to employment, a procedure acceptable only <u>faute de mieux</u>. The purification procedure tells us that too much got allocated to

Business Services; more should go to, say, Hospitals or Trade.

We have calculated a Value Added row by the purification procedure. Strictly speaking, it may be inconsistent with the detail in the other rows, so we shall later calculate a row which is the residual of product outputs less the purified inputs into them. Meanwhile, the present row bears out the Pythagorean contention of esthetic standards implicit in mathematics, for our calculations reveal that there is no Value Added in TV and Radio; it is all advertising.

```
 SLBRCUTINE PLRIFY(NCNEG)
 CCMMCN N,TCLER,            INDR(1CC),INDC(3C00),S(3C00),P(100),FLCW(
11C0), C(100)
      LET M(J,I) BE THE FRACTICN OF PRCDUCT I MADE BY ESTABLISHMENTS
  IN INCUSTRY J
   S IS THE NCN-ZERC, NCN-DIAGCNAL ELEMENTS OF M STORED ROW-BY ROW.
  INCC(I) IS THE CCLUMN NUMBER (IN M) OF THE ELEMENT S(I).
  INDR(J) IS THE NUMBER (IN S) CF THE FIRST ELEMENT OF THE JTH
      RCW CF M.
   N IS THE DIMENSICN CF M
   TCLER IS THE  IS MAXIMUM TCLERABLE DISCREPANCY
   FLCW IS THE RCW CF THE PRIMARY INPUT-OUTPUT MATRIX BEING PURIFIED
  P BECCMES THE PLRIFIEC RCW.
   NCNEG PCSITIVE GIVES NO NEGATIVE ENTRIES IN THE PURIFIED MATRIX
  ASSIGN 30 TC NEXT
  IF(NCNEG.EC.0) ASSIGN 38 TO NEXT
   INITIAL P ANC C, THE CHANGE VECTOR
  DC 15 I = 1,N
  C(I) = 0.
15 P(I) = FLCW(I)
  M IS THE ITERATICN CCUNT
  M = 0
18 DC 40 J = 1,N
   NL ANC NU ARE LCWER ANC UPPER LIMITS OF ROW J IN S.
  NL = INDR(J)
  NU = INDR(J+1) - 1
   IF CCLUMN J HAS NC ELEMENTS, SKIP ON TO NEXT COLUMN.
  IF(NL.GT.NU) GC TC 4C
   FIRST WE BRING MATERIALS INTC THE PRIMARY INDUSTRY AS IF WE
       EXPECTEC NC SHCRTAGE IN J, THE INDUSTRY WE ARE ROBBING.
   SUM BECOMES TCTAL AMCUNT CF RCBBERY.
  SUM = 0.
  DC 20 I = NL,NU
   IS = INDC(I)
  R = P(IS)*S(I)
  SUM = SUM + R
20 C(IS) = C(IS) + R
    DIC WE STEAL MCRE FRCM J THAN J HAD/  IF NCT, GC TO 38
  GC TC NEXT,(3C,38)
30 R = FLCW(J) - SUM
  IF(R.GE.0.) GC TC 38
  R = 1. - FLOW(J)/SUM
   WE SHAL HAVE TC SCALE DCWN CUR PLUNDERING BY THE FACTOR R.
  DC 35 I = NL,NU
  IS = INDC(I)
35 C(IS) = C(IS) - R*P(IS)*S(I)
  SUM = FLOW(J)
   SUBTRACT FRCM C(J) THE SUM ALLOCATED TO OTHER INDUSTRIES
38 C(J) = C(J) - SUM
40 CCNTINUE
    NEW P IS FLCW + C.  NOW COMPARE IT WITH CLD P, STORE NEW P  IN
       P, ANC CLEAR C FCR NEXT ITERATION.
   IMAX = 1
   DISMAX = 0
   DC 50 I = 1,N
```

```
      R = FLCW(I) + C(I)
      DIS = P(I) - R
      P(I) = R
      DIS = ABS(DIS)
      IF(DIS.LE.DISMAX) GO TO 50
   45 IMAX = I
      DISMAX = DIS
   50 C(I) = 0.
      WRITE(6,52) IMAX,DISMAX
   52 FORMAT(7H IMAX = I5,9H DISMAX = F10.2)
      M = M + 1
      IF(M.GE. 10) GO TO 60
      IF(DISMAX.GT.TOLER) GO TO 18
   60 RETURN
      END
```