DISAGGREGATING INPUT-OUTPUT MODELS

Alan M. Wolsky*

Abstract—A general solution is given to the problem: with little computation, construct a more disaggregated input-output model from an available model, using information about some activities that were aggregated when the available model was built. The solution is exact and explicit. When only partial information is available, the exact solution guides its user to inexpensive and precise sensitivity analysis. Detailed consideration is accorded to the solution of the most frequently occurring problem—the disaggregation of one sector into two.

I. Introduction

PUBLISHED input-output tables provide the most detailed and most comprehensive estimates available of interindustry flows. Unfortunately, they are often not detailed enough. When policy makers ask for estimates of the indirect effects, as well as the direct effects, of a change in the output of technology that has been aggregated with others, there is usually no ready answer. There is a disaggregation problem, which the work reported here will help solve.

Disaggregation is usually mentioned in work addressing aggregation, but the two are different problems. Aggregation is the concern of those who build input-output tables. Since the data needed to construct large tables are expensive to collect and corroborate, attention has been given to reasons why detailed data are not needed. Situations have been defined in which aggregation is completely harmless (Hatanka, 1952). Since these situations almost never occur, others have studied more typical cases. Theil (1957) studied the error induced by aggregation. Fisher (1958, 1961, 1966, 1969) and Kossov (1972) have sought guides to aggregation that would, in various senses, minimize the resulting errors. To the same end, Blin and Cohen (1977) have developed a hierarchy of classifications to order aggregations.

Our concern is disaggregation. We imagine that a Leontief matrix and its inverse are in the hands of a potential user whose purpose is thwarted by their high level of aggregation. The user needs a way to combine a detailed knowledge of a particular industry with the information about the rest of the economy embodied in the available matrices. Often, he must assay the sensitivity of the hoped-for answer to as yet uncollected data, particularly when time and money are limited. Few have addressed themselves to theoretical consideration of this situation. Rohn (1980), Bullard and Sebald (1977), Sebald (1974), and Tomlin (1973) have studied aspects of the sensitivity problem without regard for its relation to disaggregation. Fei (1956) has directly addressed disaggregation. He suggested that a useful and easily computed proxy for the disaggregated matrix might be found, and he described one candidate for this proxy. Fei's thought anticipates part of my own. However, I have gone beyond his goal and have found an exact, easily computed expression for the inverse of the disaggregated Leontief matrix. The results presented here can change a problem that required a computer into one that can be done by hand or with the help of a pocket calculator.[1] Just as important is the fact that, when necessary data are not available, the same results can be used for precise sensitivity analysis.

To achieve these results, we will make as much use as possible of the available but too highly aggregated model. From its Leontief matrix and few additional data, we will construct another matrix—the augmented matrix—that gives a simple (in fact, too simple) description of the disag-

Received for publication March 8, 1982. Revision accepted for publication June 14, 1983.

*Argonne National Laboratory.

This research was supported by the United States Department of Energy (DOE), under contract W-31-109-Eng-38, in order to provide improved means for evaluating the energy-conserving potential of alternate materials utilization, Jerome F. Collins was the DOE project officer.

I am grateful to J. Gibbons, B. Hannon, and W. K. Viscusi for calling pertinent references to my attention, to B. McInnis, P. Jones, and W. K. Viscusi for comments on an earlier draft, to S.-Y. Shen and J. Anderson for making it possible for me to study the distribution of values of the entries in the 1972 U.S. Input-Output Tables, and to C. Mendell for typing this paper.

[1] Let n denote the number of sectors in the aggregated model and let s denote the number of sectors that will be disaggregated into d sectors ($d > 2s$) in the disaggregated model; then the total number of sectors in the disaggregated model will be $n - s + d$. While conventional techniques for matrix inversion would require $O((n - s + d)^3)$ steps, the approach to be elaborated here requires only $O(d^3)$ steps. Since n is usually 85 or 496, while s is usually 1 and the d is usually 2 or 3, very large savings of effort can be gained.

gregated model. Then this description will be corrected by adding to the first matrix a second matrix—the distinguishing matrix—that embodies the other data needed to accurately describe the disaggregated model. The sum of these two matrices is the Leontief matrix of the disaggregated model. By regarding this matrix as a sum and by using the inverse of the Leontief matrix for the available but too highly aggregated model, we will see how the inverse of the disaggregated Leontief matrix model can be easily computed. We will also see that knowledge of the entries in the augmented matrix is sufficient for setting bounds on the entries in the distinguishing matrix. When some or all of the entries in the distinguishing matrix are not known, the bounds can suggest which of the entries are most important to estimate.

The exposition of these ideas will proceed as follows. Section II reviews the rule for aggregating sectors, presents the notion of augmented matrices, and distinguishes them from the disaggregated matrices that we seek. Section III shows how the earlier discussion yields the exact relation between the disaggregated Leontief matrix and its inverse. Section IV illustrates the use of our results with a simple example. In section V we summarize and conclude.

II. Aggregating, Augmenting, Distinguishing, and Adjusting

First, we review the aggregation of several sectors into one. Next, we present a device that enlarges (augments) an aggregated model to one that is the same size as the disaggregated model we want. We then describe the information that must be added to the enlarged (augmented) model to let us recover the disaggregated one. This information lets us distinguish sectors that were aggregated together. Finally we show how to adjust the inverse of one matrix to obtain the inverse of another. This adjustment will allow us to pass from the enlarged model to the disaggregated model. At each step, notation that will later be useful is introduced.

Consider two input-output matrices, $\mathbf{a}$ and $\mathbf{A}$, that model the same economy. Let $\mathbf{a}$ be $n \times n$ and, for simplicity of exposition, let $\mathbf{A}$ be $(n + 1) \times (n + 1)$, with the first $n - 1$ sectors of both models identical, as indicated by equation (1). (A mnemonic may be helpful: the upper case or larger letter denotes a matrix with more sectors than one denoted by a lower case or smaller letter.) Let us recall the familiar relation, equations (2), (3), (4) below, between sectors and an aggregate formed from them. (Throughout this section, $i, j = 1, \ldots, n - 1$, and in the rest of this paper $i, j = 1, \ldots, n$.)

$$\mathbf{a}_{ij} = \mathbf{A}_{ij} \tag{1}$$

$$\mathbf{a}_{nj} = \mathbf{A}_{nj} + \mathbf{A}_{n+1,j} \tag{2}$$

$$\mathbf{a}_{in} = w_1\mathbf{A}_{in} + w_2\mathbf{A}_{i,n+1} \tag{3}$$

$$\mathbf{a}_{nn} = w_1(\mathbf{A}_{nn} + \mathbf{A}_{n+1,n}) + w_2(\mathbf{A}_{n,n+1} + \mathbf{A}_{n+1,n+1}). \tag{4}$$

The parameter w_1 (respectively, w_2) is the ratio of the gross output of the n^{th} sector (respectively $(n + 1)^{\text{th}}$ sector) to the sum of the gross outputs of the n^{th} and $(n + 1)^{\text{th}}$ sectors. When more than two sectors are aggregated, the aggregated and disaggregated models are related by the straightforward generalization of the equations just given.

Since our goal is to disaggregate, not aggregate, one might try to invert these formulas to obtain $\mathbf{A}$ from $\mathbf{a}$. Of course, this is impossible because $\mathbf{A}$ embodies more information than $\mathbf{a}$ and thus $\mathbf{a}$ does not uniquely define $\mathbf{A}$. We accept this and temporarily put aside our effort to characterize the disaggregated model. Instead we now focus our attention on one matrix, A, from among the many $(n + 1) \times (n + 1)$ input-output matrices that satisfy equations (1)–(4). The matrix A will serve as a convenient half-way house on our trip from the aggregated model, described by $\mathbf{a}$, to the disaggregated model described by $\mathbf{A}$. (An upper case or large letter is used to denote the matrix that will be defined because its dimension is the same as the dimension of the disaggregated matrix. Please note that the identity matrix, no matter what its dimension, will be denoted by I.) The convenience of A comes from the ease with which its Leontief matrix, $L \equiv I - A$, can be inverted when one knows the inverse of the aggregated model's Leontief matrix, $\mathbf{l} \equiv I - \mathbf{a}$. Later, it will be shown that this characteristic of A reduces the burden of inverting the Leontief matrix, $\mathbf{L} \equiv I - \mathbf{A}$, of the disaggregated model.

To construct A, imagine that $\mathbf{a}$ arose from the aggregation of two essentially identical sectors. More specifically, let the first $n - 1$ sectors described by A be the sectors that are common to $\mathbf{a}$

and $\mathbf{A}$, and let the n^{th} and $(n+1)^{\text{th}}$ sectors described by A have technologies identical to the technology described by the n^{th} column of $\mathbf{a}$. Furthermore, let the output of the n^{th} sector and the output of the $(n+1)^{\text{th}}$ sector be supplied to each of the other sectors in fixed proportion to the output of the aggregate. These instructions are made explicit in the definition of A given next.

$$A_{ij} \equiv \mathbf{a}_{ij} \tag{5}$$

$$A_{nj} \equiv w_1 \mathbf{a}_{nj} \text{ and } A_{n+1,j} \equiv w_2 \mathbf{a}_{nj} \tag{6}$$

$$A_{in} = A_{i,n+1} \equiv \mathbf{a}_{in} \tag{7}$$

$$\begin{pmatrix} A_{nn} & A_{n,n+1} \\ A_{n+1,n} & A_{n+1,n+1} \end{pmatrix} \equiv \mathbf{a}_{nn} \begin{pmatrix} w_1 & w_1 \\ w_2 & w_2 \end{pmatrix}. \tag{8}$$

With this definition of A, L^{-1} can be easily obtained from $\mathbf{l}^{-1}$.

$$L_{ij}^{-1} = \mathbf{l}_{ij}^{-1} \qquad L_{nj}^{-1} = w_1 \mathbf{l}_{nj}^{-1} \tag{9}$$

and

$$L_{n+1,j}^{-1} = w_2 \mathbf{l}_{nj}^{-1} \tag{10}$$

$$L_{in}^{-1} = L_{i,n+1}^{-1} = \mathbf{l}_{in}^{-1} \tag{11}$$

$$\begin{pmatrix} L_{nn}^{-1} & L_{n,n+1}^{-1} \\ L_{n+1,n}^{-1} & L_{n+1,n+1}^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (\mathbf{l}_{nn}^{-1} - 1) \begin{pmatrix} w_1 & w_1 \\ w_2 & w_2 \end{pmatrix}. \tag{12}$$

We see that L and L^{-1} are constructed from $\mathbf{l}$ and $\mathbf{l}^{-1}$ and the additional information embodied in w_1 and $w_2 = 1 - w_1$. Fortunately, these parameters are relatively easy to estimate. They are the ratios of gross outputs, and these gross outputs can be inferred from the financial statements of firms or from statistics published by trade associations.

When one sector is to be disaggregated into more than two sectors ($s = 1$, $d > 2$), the appropriate A is constructed like the one described in equations (5)–(8). The pertinent column of $\mathbf{a}$ is replaced by as many identical columns as there are sectors to be created. The matching row of $\mathbf{a}$ is replaced by the same number of rows, each of which is the product of a weighting factor and the original row of A. The inverse, L^{-1}, of the associated Leontief matrix is again found from $\mathbf{l}^{-1}$ by formulas like equations (9)–(12). When several sectors must be disaggregated ($s > 1$, $d \geq 2s$), the prescription just given should be applied to each sector. In that case, a convention will expedite analysis; one relabels the sectors to be disaggregated so that they are subscripted with the highest numbers. This procedure will let us later use a result concerning partitioned matrices.

To disaggregate, we must consider more information than is embodied in the augmented matrix. Since this matrix describes the newly created sectors as being essentially the same, we must distinguish them from each other. Thus we introduce a new matrix, denoted by Δ and called the distinguishing matrix.

$$\Delta \equiv \mathbf{A} - A = L - \mathbf{L}. \tag{13}$$

Equations (2)–(4) imply that when one sector is disaggregated into two, the matrix elements of the distinguishing matrix can be parameterized by the independent variables δ_i, δ_n, σ_j, σ_n, and ξ, as shown next:[2]

$$\Delta_{in} = w_2 \delta_i \quad \text{and} \quad \Delta_{i,n+1} = -w_1 \delta_i \tag{14}$$

$$\Delta_{nj} = \sigma_j = -\Delta_{n+1,j} \tag{15}$$

$$\begin{pmatrix} \Delta_{nn} & \Delta_{n,n+1} \\ \Delta_{n+1,n} & \Delta_{n+1,n+1} \end{pmatrix} = \frac{1}{2}\delta_n \begin{pmatrix} w_2 & -w_1 \\ w_2 & -w_1 \end{pmatrix} + \sigma_n \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} + \xi \begin{pmatrix} w_2 & -w_1 \\ -w_2 & w_1 \end{pmatrix}. \tag{16}$$

The independent variables have the following meanings: δ_i manifests the difference between the n^{th} and $(n+1)^{\text{th}}$ sectors in their demand for input from the i^{th} sector; σ_j manifests the departure from average in what the n^{th} and $(n+1)^{\text{th}}$ sectors supply to the j^{th} sector; δ_n, σ_n, and ξ manifest like quantities for intra-aggregate exchanges. When the distinguishing matrix, Δ, is unknown or incompletely known, it is helpful to know that although δ_i, δ_n, σ_j, σ_n, and ξ are defined as above, these parameters can be bounded using knowledge of only w_1, w_2, and $\mathbf{a}$, as explained in appendix A.

Using these ideas, the problem of disaggregating published input-output tables—or, what is the same, using them in conjunction with knowledge of the constituents of one or more of their sectors—is overcome by accomplishing two tasks. First, one must estimate the weighting factor of each of the newly distinct sectors and the distinguishing matrix. With this done, $\mathbf{L} = I - \mathbf{A}$ can be expressed in the form $\mathbf{L} = L - \Delta$. That this ex-

[2] Although there are $4n$ entries in this distinguishing matrix that are not identically zero, there are only $2n + 1$ independent variables.

pedites the second task, the inversion of **L**, will be shown in the next section.

To prepare for that demonstration, another idea will be introduced. It will be shown how the inverse of a matrix can be adjusted so that the result is the inverse of another matrix. Consider two nonsingular matrices of the same dimension, M_I and M_{II}, and their difference D. The inverse of M_{II} can be obtained from D and the inverse of M_I by the following manipulation:

$$M_{II}^{-1} = (M_I - D)^{-1} = \left[M_I(I - M_I^{-1}D)\right]^{-1} \tag{17}$$

$$\begin{aligned} M_{II}^{-1} &= \left(I - M_I^{-1}D\right)^{-1}M_I^{-1} \\ &= M_I^{-1} + M_I^{-1}D(I - M_ID)^{-1}M_I^{-1}. \end{aligned} \tag{18}$$

We see that $(I - M_I^{-1}D)^{-1}$ can "adjust" the inverse of M_I so as to yield the inverse of M_{II}. The usefulness of this remark stands or falls on the ease with which $(I - M_I^{-1}D)^{-1}$ can be found. The next section shows that, in cases of interest to us, finding it is not a burden.

III. Exact Disaggregation

The means to construct the inverse of the disaggregated Leontief matrix are now at hand. The construction is sketched below, and afterward we discuss the most important example of its use—the disaggregation of one sector into two sectors.

The inverse of the disaggregated Leontief matrix can be expressed in terms of the distinguishing matrix and the inverse of the augmented matrix by using the argument we developed in the previous section. Recall the right hand side of equation (18) to see one form this expression takes:

$$\mathbf{L}^{-1} = L^{-1} + L^{-1}\Delta(I - L^{-1}\Delta)^{-1}L^{-1}. \tag{19}$$

The second term adjusts the inverse of the augmented matrix so that the sum provides an exact expression for the inverse of the disaggregated Leontief matrix. As it stands, this result is compact and explicit. However, the reader may doubt its practical use. After all, the factor $(I - L^{-1}\Delta)^{-1}$ appears in the second term, and so a large matrix must still be inverted. Typical input-output matrix sizes are 85×85 and 496×496, and the inversion of either usually requires a computer. Remarkably, the reader's possible doubt is baseless. For problems that are likely to occur in practice, the computation can be done by hand or with a pocket calculator.

To see why this claim is true, one needs an analytic method of inversion; partitioning and subsequent inversion provides it (see appendix B). After partitioning L and Δ as described in section II and then forming their product, one learns that the upper left-hand block (whose dimension is $n - s$) of $L^{-1}\Delta$ is zero because of equations (1)–(4). Thus the upper left-hand block of $(I - L^{-1}\Delta)$ is the identity. This is a crucial simplification because inversion by partitioning requires the computation of two inverses: the inverse of the upper left-hand block (now seen to be the identity) and the inverse of a matrix, Γ, which has the same dimension as the lower right-hand block. This dimension is d, the total number of sectors into which the aggregated sectors are split ($d \geq 2s$).

Most often we wish to disaggregate one sector into two, and thus only a 2×2 matrix must be inverted. We discuss this case now. Three quantities must be introduced. The first, $\det\Gamma$, is the determinant of the matrix related to the lower right-hand block of $I - L^{-1}\Delta$ (see appendix B).

$$\det\Gamma = 1 + \tfrac{1}{2}(w_1 - w_2)\delta_n - \xi - \gamma \tag{20}$$

where

$$\gamma \equiv \sum_{k=1}^{n}\sum_{q=1}^{n}\sigma_k \mathbf{l}_{kq}^{-1}\delta_q.$$

The second quantity, D_i, can be thought of as the difference between gross outputs in the aggregated model that result from the difference, $(\delta_1, \ldots, \delta_n)$, between two bills of final demand:

$$D_i \equiv \sum_{q=1}^{n}\mathbf{l}_{iq}^{-1}\delta_q. \tag{21}$$

The third quantity, S_j, is a weighted sum of the gross outputs necessary, according to the aggregate model, to supply a unit of final demand from the j^{th} sector:

$$S_j \equiv \sum_{k=1}^{n}\sigma_k\mathbf{l}_{kj}^{-1}. \tag{22}$$

Although γ, D_i, and S_j are the sums of many terms, most of these terms are negligible because, in practice, most of the entries in $\mathbf{l}^{-1}$ are zero or small, as are most values of δ_i and σ_j (see appendix A for bounds).[3]

When we disaggregate a sector in a published input-output model into two distinct sectors, each sector in the resulting model is either common to both models or is one of the newly distinct sectors, and all pairs of sectors in the disaggregated model lie in one of four groups: common-distinct, distinct-common, common-common, or distinct-distinct. For each pair, we want to know how the gross output of the first sector depends on the final demand for the output of the second sector. Equation (19) provides the following answers. We begin with the dependence of the common sectors on the newly distinct ones. (Our index notation follows that in section II.)

$$\mathbf{L}_{in}^{-1} = \mathbf{l}_{in}^{-1} + (\det\Gamma)^{-1} D_i (w_2 + S_n) \quad (23)$$

$$\mathbf{L}_{i,n+1}^{-1} = \mathbf{l}_{in}^{-1} + (\det\Gamma)^{-1} D_i (-w_1 + S_n). \quad (24)$$

Now consider the response of the gross output of each of the two newly distinct sectors to final demand for output of each of the sectors common to both models.

$$\mathbf{L}_{nj}^{-1} = w_1 \mathbf{l}_{nj}^{-1} + (\det\Gamma)^{-1}(1 + w_1 D_n) S_j \quad (25)$$

$$\mathbf{L}_{n+1,j}^{-1} = w_2 \mathbf{l}_{nj}^{-1} + (\det\Gamma)^{-1}(-1 + w_2 D_n) S_j. \quad (26)$$

The third kind of dependence—that between two sectors neither of which has been disaggregated—is also interesting. Equation (19) implies that

$$\mathbf{L}_{ij}^{-1} = \mathbf{l}_{ij}^{-1} + (\det\Gamma)^{-1} D_i S_j. \quad (27)$$

The necessary adjustment is proportional to the product of two factors, each relating one of the common sectors to the distinction between the disaggregated ones. Thus to first order, no adjustment is necessary.

The response of the gross output of one of the newly distinct sectors to final demand for its own output, or to the final demand for the output of its partner, also follows from equation (19). I omit expressions for $\mathbf{L}_{nn}$, $\mathbf{L}_{n,n+1}$, $\mathbf{L}_{n+1,n}$, and $\mathbf{L}_{n+1,n+1}$ in terms of $\mathbf{l}$ and Δ because of the length of these expressions and because they are less likely to be needed, in practice, than the ones given earlier. An example of a problem in which they would be needed arises if one wanted to estimate the distinct gross outputs of organic (i.e., carbon-containing) chemicals and inorganic (i.e., not carbon-containing) chemicals in response to distinct final demands for these commodities, because present U.S. Bureau of Economic Analysis (BEA) practice aggregates all chemicals in a single sector.

Equations (23)–(27) make the result embodied in equation (19) immediately applicable to the case most often encountered in practice, the disaggregation of one sector into two. When w_1, w_2, and Δ are known, only the arithmetic remains to be done, and it can be done by hand. When the weighting factors or the distinguishing matrix is not known, the results presented in this section can be used to determine the sensitivity of the desired answer to the relevant uncertainty.

IV. An Illustration

This section illustrates the use of the results presented in other parts of this paper. To emphasize these results, we will avoid numerical detail by considering a hypothetical three sector input-output model ($n = 3$), and then we will describe the process by which one of these sectors ($s = 1$) is disaggregated into two sectors ($d = 2$) so as to provide a more disaggregated model with four sectors ($n - s + d = 4$).

Let us suppose that the model specified by equations (28) and (29) is available, perhaps from a government agency or the collaboration of several investigators, but that this model is too highly aggregated to be immediately useful.

$$\mathbf{a} \equiv \begin{pmatrix} 0.02 & 0.05 & 0.05 \\ 0.35 & 0.10 & 0.45 \\ 0.08 & 0.25 & 0.15 \end{pmatrix} \quad (28)$$

$$\mathbf{l}^{-1} = (I - \mathbf{a})^{-1} = \begin{pmatrix} 1.06 & 0.09 & 0.11 \\ 0.54 & 1.35 & 0.75 \\ 0.26 & 0.41 & 1.41 \end{pmatrix}. \quad (29)$$

It might be necessary to estimate the impact of proposed government regulation or deregulation that would cause one commodity to substitute for another when both commodities are products of

[3] The most recent U.S. input-output tables (Ritz, 1979) illustrate our remark about the great number of small entries in $\mathbf{l}^{-1}$. Although the largest entry in the 496 × 496 table (called the commodity by commodity Total Requirements Matrix by the U.S. BEA) is 1.590, 92% of the entries are less than or equal to 0.005 and 95% of the entries are less than or equal to 0.010. A similar situation exists at the 85 × 85 level of aggregation.

the same sector in the available model. We will assume that we have to estimate such an impact. For example, one might need to distinguish naturally sweetened from artificially sweetened soft drinks, both now aggregated by the U.S. BEA in the sector Bottled and Canned Soft Drinks. From equation (27) we learned that disaggregating sectors other than the one in question will affect our numerical answers in second order and so, in practice, one would begin by disaggregating only the sector in question. This is what we do in this illustration.

In accordance with equations (5)–(8) and (13)–(16), the disaggregated input-output matrix has the following form:

$$\mathbf{A} = \begin{pmatrix} 0.02 & 0.05 & 0.05 & 0.05 \\ 0.35 & 0.10 & 0.45 & 0.45 \\ w_1 \times 0.08 & w_1 \times 0.25 & w_1 \times 0.15 & w_1 \times 0.15 \\ w_2 \times 0.08 & w_2 \times 0.25 & w_2 \times 0.15 & w_2 \times 0.15 \end{pmatrix} + \begin{pmatrix} 0 & 0 & w_2\delta_1 & -w_1\delta_1 \\ 0 & 0 & w_2\delta_2 & -w_1\delta_2 \\ \sigma_1 & \sigma_2 & \left(\tfrac{1}{2}\delta_3 + \xi\right)w_2 + \sigma_3 & -\left(\tfrac{1}{2}\delta_3 + \xi\right)w_1 + \sigma_3 \\ -\sigma_1 & -\sigma_2 & \left(\tfrac{1}{2}\delta_3 - \xi\right)w_2 - \sigma_3 & -\left(\tfrac{1}{2}\delta_3 - \xi\right)w_1 - \sigma_3 \end{pmatrix}. \quad (30)$$

The first term on the right-hand side is the augmented matrix and the second term is the distinguishing matrix.

Our first step is to estimate, for the year on which the available model is based, the ratio of the gross output of one of the newly distinct sectors to the gross output of the aggregate in which that sector had been lumped. That is to say, we must estimate either w_1 or w_2; recall that $w_1 + w_2 = 1$. In practice, it is best to consider both ratios and to choose the one that can be most reliably estimated with the desired effort. Let us suppose that data published by a suitable trade association allow us to infer that $w_1 = 0.75$ and $w_2 = 0.25$.

We also want estimates for the other unknown parameters, δ_i, σ_j, and ξ. Since reliable estimates are much more difficult to obtain for them than for w_1 or w_2, we do not immediately seek those estimates. Instead, our second step is to bound the unknown parameters by using the results in appendix A and the values of w_1, w_2, and **a**; one would later regret the effort to estimate the unknown parameters if they could have been judged negligible at this stage. For the three-sector model given above and the values $w_1 = 0.75$ and $w_2 = 0.25$, equations (A.1), (A.2), (A.7), (A.8), and (A.9) have been used to obtain the following bounds:

$$-0.20 = \max\left\{-\frac{0.05}{0.25}, -\frac{0.95}{0.75}\right\} \le \delta_1 \le \min\left\{\frac{0.05}{0.75}, \frac{0.95}{0.25}\right\} = 0.07 \quad (31)$$

$$-0.73 = \max\left\{-\frac{0.45}{0.25}, -\frac{0.55}{0.75}\right\} \le \delta_2 \le \min\left\{\frac{0.45}{0.75}, \frac{0.55}{0.25}\right\} = 0.60 \quad (32)$$

$$-0.60 = \max\left\{-\frac{0.15}{0.25}, -\frac{1.85}{0.75}\right\} \le \delta_3 \le \min\left\{\frac{0.15}{0.75}, \frac{1.85}{0.25}\right\} = 0.20 \quad (33)$$

$$-0.06 \le \sigma_1 \le 0.02 \quad (34)$$

$$-0.19 \le \sigma_2 \le 0.06 \quad (35)$$

$$-0.11 \le \sigma_3 \le 0.04 \quad (36)$$

$$-0.29 \le \xi \le 0.29. \quad (37)$$

(The computation of the bounds given by equations (34)–(37) is omitted because of its length.) Although we would hesitate before neglecting any of the unknown parameters in the illustration, we certainly would defer effort to estimate σ_1, σ_2, and δ_1 in favor of effort to estimate δ_2. On the other hand, when sufficient resources are available, the bounds described in appendix A can serve as checks on the consistency of the available model and the data that might be developed in the course of the project.

The next step in any disaggregation project is the collection of pertinent data. This is taxing, and no amount of mathematics can let us completely evade it. Let us suppose that our efforts resulted in estimates $\delta_1 = -0.10$, $\delta_2 = 0.30$, $\delta_3 = -0.40$, $\sigma_1 = -0.06$, $\sigma_2 = 0.03$, $\sigma_3 = -0.10$ and $\xi = 0.20$. (Note that although the bounds on δ_2 may be wider than those on δ_3, nothing prevents the absolute value of δ_3 from exceeding that of δ_2.) The values given here and those for w_1 and w_2, together with the values given for the available input-output matrix in equation (28), specify the disaggregated input-output matrix.[4]

The last step is the computation of the exact values of the entries in the inverse of the disaggregated Leontief matrix. To obtain these entries we must compute det Γ according to equation (20), which for our illustration yields:

$$\det \Gamma = 1 - 0.36 = 0.64. \tag{38}$$

Now we can compute any of the entries in the inverse of the disaggregated Leontief matrix according to equations (23)–(27). For example, we calculate the gross outputs of the common sectors, the first and second, generated by production in the newly distinguished sectors, the third and fourth, to satisfy final demand. According to equations (23)–(24), one obtains the following values:

$$\mathbf{L}_{13}^{-1} = 0.11 + (0.64)^{-1} \times (-0.12) \times (0.25 - 0.12) = 0.09 \tag{39}$$

$$\mathbf{L}_{14}^{-1} = 0.11 + (0.64)^{-1} \times (-0.12) \times (-0.75 - 0.12) = 0.27. \tag{40}$$

Early in this section, we remarked that for a project concerning the substitution of one commodity for another from the same sector, one would begin by disaggregating that sector, and this is how our simple illustration was handled. As the project unfolds, we must decide whether this disaggregation ($s = 1$, $d = 2$) is sufficient for our purposes. If we divide soft drinks into naturally and artificially sweetened soft drinks, we must later ask if the "chemicals" sector should then be disaggregated into the "saccharin" sector and the "all other chemicals" sector. To answer, we would go through the same steps, except one, as we have in this example. The exception is the step in which we gather data that yields reasonably definite values for δ_i, σ_j, and ξ. In this case, we would only seek their orders of magnitude. Then we would use equation (27) to indicate whether the disaggregation of the chemicals sector would significantly affect our estimates for soft drinks.

[4]The disaggregated input-output matrix specified (within rounding errors) by the hypothetical data in our example is

$$\mathbf{A} = \begin{pmatrix} 0.02 & 0.05 & 0.03 & 0.13 \\ 0.35 & 0.10 & 0.53 & 0.23 \\ 0.00 & 0.22 & 0.01 & 0.01 \\ 0.08 & 0.03 & 0.04 & 0.44 \end{pmatrix}.$$

By construction, the matrix presented in equation (35) is obtained from the above after its third and fourth sectors are aggregated. Since the disaggregated model of our example has only four sectors, one can calculate the inverse of the model's disaggregated Leontief matrix directly from **A** by hand or pocket calculator. The result (within rounding error) is

$$\mathbf{L} \equiv (I - \mathbf{A})^{-1} = \begin{pmatrix} 1.08 & 0.09 & 0.09 & 0.28 \\ 0.54 & 1.35 & 0.76 & 0.68 \\ 0.12 & 0.30 & 1.18 & 0.17 \\ 0.19 & 0.11 & 0.13 & 1.87 \end{pmatrix}.$$

V. Summary and Conclusion

The results reported here show how an input-output model can be disaggregated to describe subsectors of some of its sectors. This process can achieve as great a degree of disaggregation as the user has the data to support or wishes to perform for sensitivity analyses. The starting model's Leontief matrix and its inverse are used to render the subsequent disaggregation analytically transparent and computationally simple. The general solution is presented at the beginning of section III.

Throughout this paper, emphasis is given to the problem most often encountered in practice—disaggregating one sector into two. When the necessary data have already been collected, formulas in section III embody the exact solution. The same formulas are the basis for accurate sensitivity analysis when pertinent data are uncertain. The associated computation can be done on a pocket calculator or by hand. In addition, appendix A presents inequalities that can often reduce the uncertainty surrounding unknown data.

When one or more sectors must be disaggregated into many sectors (e.g., $d \geq 10$), hand computation is no longer practical. The work presented here is still pertinent because the general solution can provide the basis for an algorithm that will let us use microcomputers to accomplish disaggregations that can now only be done using larger machines. Nonetheless, it is unusual to have enough data to make extensive disaggregation worthwhile. Most often one has detailed knowledge, or a detailed hypothesis, about the constituents of only one sector in an available input-output model, and one has the need to estimate various indirect effects of intrasectoral changes. Such is often the case when one attempts to use regional input-output tables, usually highly aggregated, or the tables, also highly aggregated, that describe less developed countries.[5] It is also the

[5]I thank an anonymous referee for calling my attention to this.

case when one attempts to discuss the effects of specific subsidies or regulations on the U.S. economy. It is in such discussions that our derivations can be put to immediate use.

APPENDIX A

Bounds on the Elements of the Distinguishing Matrix

Section III shows that the inverse of the disaggregated Leontief matrix depends on the distinguishing matrix. When the distinguishing matrix embodies data that are unknown or incompletely known, our immediate recourse is to sensitivity analysis. When recourse to sensitivity analysis is necessary, two considerations are pertinent. First, the entries of the distinguishing matrix are not independent of each other (e.g., see equations (14)–(16)) because of equations (1)–(4). Second, even after the dependence of the entries has been taken into account, the remaining free variables can be bounded using the weighting factors and the entries in the aggregated input-output matrix because these entries result from the addition of nonnegative numbers. For example, if $\mathbf{a}_{nj}$ equals zero, then both $\mathbf{A}_{nj}$ and $\mathbf{A}_{n+1,j}$ must equal zero. More detailed consideration gives stronger results, as we will now illustrate.

In the case of one sector disaggregated into two, parameterizing the distinguishing matrix in terms of δ_i, σ_j, and ξ, according to equations (14)–(16), takes full account of the dependence implied by equations (1)–(4). However, more can be inferred. The facts that $\mathbf{a}_{in}$ is a convex combination of $\mathbf{A}_{in}$ and $\mathbf{A}_{i,n+1}$ and that these numbers are nonnegative imply that

$$\max\{-w_2^{-1}\mathbf{a}_{in}, -w_1^{-1}(1-\mathbf{a}_{in})\} \le \delta_i \le \min\{w_1^{-1}\mathbf{a}_{in}, w_2^{-1}(1-\mathbf{a}_{in})\}. \quad (A.1)$$

Because $\mathbf{a}_{nj}$ is the sum of $\mathbf{A}_{nj}$ and $\mathbf{A}_{n+1,j}$, both of which are nonnegative, it is true that

$$\max\{-w_1\mathbf{a}_{nj}, -1+w_2\mathbf{a}_{nj}\} \le \sigma_j \le \min\{w_2\mathbf{a}_{nj}, 1-w_1\mathbf{a}_{nj}\}. \quad (A.2)$$

Likewise, since $\mathbf{a}_{nn}$ is a weighted sum of $\mathbf{A}_{nn}$, $\mathbf{A}_{n,n+1}$, $\mathbf{A}_{n+1,n}$, and $\mathbf{A}_{n+1,n+1}$ and since each of these is nonnegative, the following four compatible linear inequalities hold:

$$-w_1\mathbf{a}_{nn} \le +\tfrac{1}{2}w_2\delta_n + \sigma_n + w_2\xi \le 1 - w_1\mathbf{a}_{nn} \quad (A.3)$$
$$-w_1\mathbf{a}_{nn} \le -\tfrac{1}{2}w_1\delta_n + \sigma_n - w_1\xi \le 1 - w_1\mathbf{a}_{nn} \quad (A.4)$$
$$-w_2\mathbf{a}_{nn} \le +\tfrac{1}{2}w_2\delta_n - \sigma_n - w_2\xi \le 1 - w_2\mathbf{a}_{nn} \quad (A.5)$$
$$-w_2\mathbf{a}_{nn} \le -\tfrac{1}{2}w_1\delta_n - \sigma_n + w_1\xi \le 1 - w_2\mathbf{a}_{nn}. \quad (A.6)$$

These inequalities can be "solved" to yield three inequalities that bound each variable without reference to the others.

$$\max\{-w_2^{-1}\mathbf{a}_{nn}, -w_1^{-1}(2-\mathbf{a}_{nn})\} \equiv \underline{\delta}_n \le \delta_n \le \bar{\delta}_n \equiv \min\{w_1^{-1}\mathbf{a}_{nn}, w_2^{-1}(2-\mathbf{a}_{nn})\} \quad (A.7)$$
$$\max\{-w_1\mathbf{a}_{nn}, -1+w_2\mathbf{a}_{nn}\} \equiv \underline{\sigma}_n \le \sigma_n \le \bar{\sigma}_n \equiv \min\{w_2\mathbf{a}_{nn}, 1-w_1\mathbf{a}_{nn}\} \quad (A.8)$$
$$\underline{\xi} \le \xi \le \bar{\xi} \quad (A.9)$$

where

$$\bar{\xi} \equiv (2w_1w_2)^{-1}\{\min w_1 + (w_1 - 2w_1^2)\mathbf{a}_{nn} - 2w_1\underline{\sigma}_n, w_2 + (w_2 - 2w_2^2)\mathbf{a}_{nn} + 2w_2\bar{\sigma}_n\} \quad (A.10)$$
$$\underline{\xi} \equiv (2w_1w_2)^{-1}\{\max -w_1 + (w_1 - 2w_1^2)\mathbf{a}_{nn} - 2w_1\bar{\sigma}_n, -w_2 + (w_2 - 2w_2^2)\mathbf{a}_{nn} + 2w_2\underline{\sigma}_n\}. \quad (A.11)$$

When time or money limit the effort that can be expended estimating the entries in the distinguishing matrix, the inequalities presented here may help decide whether these entries are too small to pursue further.

APPENDIX B

Exact Inversion by Partitioning

Recall that if M is a nonsingular matrix, its inverse can be obtained by the scheme indicated below:

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad (B.1)$$

$$M^{-1} = \begin{pmatrix} M_{11}^{-1} + M_{11}^{-1}M_{12}\Gamma^{-1}M_{21}M_{11}^{-1} & -M_{11}^{-1}M_{12}\Gamma^{-1} \\ -\Gamma^{-1}M_{21}M_{11}^{-1} & \Gamma^{-1} \end{pmatrix} \quad (B.2)$$

$$\Gamma \equiv M_{22} - M_{21}M_{11}^{-1}M_{12}. \quad (B.3)$$

In general, both M_{11} and Γ must be inverted. In the case of interest to us, $M = I - L^{-1}\Delta$, and in this case $M_{11} = I$ because the definitions of L and Δ, which were motivated by equations (1)–(4), entail the equality $[L^{-1}\Delta]_{11} = 0$.

$$M_{11} = I,\ M_{12} = -[L^{-1}\Delta]_{12},\ M_{21} = -[L^{-1}\Delta]_{21}, \quad M_{22} = I - [L^{-1}\Delta]_{22} \quad (B.4)$$

and

$$\Gamma \equiv I - [L^{-1}\Delta]_{22} - [L^{-1}\Delta]_{21}[L^{-1}\Delta]_{12}. \quad (B.5)$$

Now we need only compute the inverse of the square matrix, Γ, whose number of rows (and columns) is just the number of newly distinguished sectors.

REFERENCES

Blin, Jean-Marie, and Claude Cohen, "Technological Similarity and Aggregation in Input-Output Systems: A Cluster-Analytic Approach," this REVIEW 49 (Feb. 1977), 82–91.

Bullard, Clark W., and Anthony V. Sebald, "Effects of Parametric Uncertainty and Technological Change on Input-Output Models," this REVIEW 49 (Feb. 1977), 75–81.

Fei, John Ching-Han, "A Fundamental Theorem for the Aggregation Problem of Input-Output Analysis," *Econometrica* 24 (1956), 400–412.

Fisher, Walter D., "Criteria for Aggregation in Input-Output Analysis," this REVIEW 40 (Aug. 1958), 250–260.

______, "Optimal Aggregation in Multi-Equation Prediction Models," *Econometrica* 30 (1961), 764–769.

______, "Simplification in Economic Models," *Econometrica* 34 (1966), 563–584.

______, Clustering and Aggregation in Economics (Baltimore: The Johns Hopkins Press, 1969).

Hatanka, M., "Note on Consolidation within a Leontief System," *Econometrica* 20 (Apr. 1952), 301–303.

Kossov, Vassily, "The Theory of Aggregation in Input-Output Models," in A. P. Carter and A. Brody (eds.), *Contributions to Input-Output Analysis* (New York: North-Holland Publishing Co., 1972).

Ritz, Philip M., *The Detailed Input-Output Structure of the U.S. Economy: 1972* (Washington, D.C.: Superintendent of Documents, U.S. G.P.O. stock number 003-010-00064-3 and 65-1, 1979).

Rohn, Jiri, "Input-Output Model with Interval Data," *Econometrica* 48 (Apr. 1980), 767–769.

Sebald, Anthony, "An Analysis of the Sensitivity of Large Scale Input-Output Models to Parametric Uncertainty," Center for Advanced Computation, document no. 122, University of Illinois at Urbana, June 1974.

Theil, Henri, "Linear Aggregation in Input-Output Analysis," *Econometrica* 25 (Jan. 1957), 111–127.

Tomlin, Paul, "Input-Output Error Analysis—Present Status," U.S. Census Bureau, research documentation no. 73-311, 1973.