

A model of the type used to forecast the effects on an industry of changes in the national economy is described. The components representing input data and the type of output from such a model are discussed.

This paper is intended as a tutorial discussion of an advanced application in data processing.

An input-output econometric model

by K. S. Sarma

Input-output models are used to forecast various effects that can occur to an industry as it interacts with other industries under changing conditions in the general economy. The conceptual framework for such models was established many years ago by such economic theoreticians as Quesnay (1758) and Leon Walras (1877) and, more recently, Leontief. As is readily apparent, these interindustry interactions are not only complex but numerous. Therefore, the practical use of these concepts has largely awaited the advent of the computer. In this paper, an input-output model developed for IBM is described to show what parameters are used in such models and what type of results are expected.

The purpose of the IBM input-output (I/O) model is to forecast the industrial implications of changes in the overall national economy. Given forecasts of such broad components of the Gross National Product (GNP) as personal consumption expenditures, investment, exports, imports, etc., the I/O model makes predictions of output originating in different industries.

The GNP components are forecast by an Annual Macro Econometric Model developed by IBM's Economic Research Department.¹ The Annual Model consists of a set of simultaneous equations which represent the interrelationships among different aggregate (macro) economic variables. From a set of assumptions about policy variables such as government expenditures, personal income tax rates, treasury bill rates, etc., the system of equations can be solved for such economic aggregates as GNP, personal consumption expenditures, investment expenditures, imports, etc. The Annual Model is thus suitable for forecasting

the impact of alternative fiscal and monetary policies, political developments, demographic trends, etc. on the overall national economy. But it cannot predict the impact of these aggregate economic changes on individual industries. The IBM I/O model, however, is specifically designed to assess these direct and indirect industrial impacts.

For example, the macro model can be used for predicting the increase in total consumer expenditures on durable goods as a consequence of a tax cut. However, the model is not capable of forecasting the resultant expansion in the output of particular consumer durable goods such as furniture, household appliances, and television sets. Also, the Annual Model cannot appraise the indirect effects of the consumer expenditure on those industries that provide raw materials to make the durable goods. The IBM I/O model, however, can assess the effect on various industries of a change in consumer spending for durables, to continue the example, by answering the following types of questions:

- What proportions of the increased consumer expenditure on durable goods, whether it resulted from a tax cut or another reason, will be for specific commodities such as household appliances, furniture and fixtures, and television sets?
- By how much will the output of the industries that manufacture the consumer durables be increased?
- By how much must the output of a raw material like steel be increased to produce the additional final products?
- By how much should coal output be increased to produce this additional amount of steel?

The list of questions would continue, but a more important question is:

- By how much will the *Gross Product Originating*, or GPO, (which is the sum of wages, salaries, profits, etc.) increase in such industries as appliance and furniture manufacturing, which are directly affected by the increase in demand for consumers' durable goods, and in industries like coal, steel, plastics, and electric utilities, which are indirectly affected by the increased demand?

The analysis of the impact of a change in consumer spending is only one example of the changes in the economy dealt with by the I/O model.

Similarly, the I/O model enables us to assess the change in industrial composition of total output of the economy as a consequence of changes in various other variables such as government expenditures, exports, or the demographic structure of the population.

The relationship between GNP components and such industry variables as final demand, output, and value added can be seen from a set of identities known as the I/O transactions tables which show the flows of goods and services among different industries as well as flows from the industries to final users.

The first section of the paper discusses the rudiments of the I/O transactions tables. The discussion is confined to a simple example, but it covers the most essential features of an I/O framework. A comprehensive discussion of the problems in constructing I/O tables, such as the treatment of imports and secondary products, can be found in Reference 2.

The transactions tables constitute the foundation upon which all I/O models are built. The second section shows how the two key sets of parameters that are used in I/O forecasting are obtained from an I/O transactions table. These parameters, the *Input-Output Coefficients* and the *Final Demand Coefficients* (also known as the *Bridge Coefficients*), are often referred to as the **A** and the **H** matrices. The I/O coefficients indicate how much output of one industry is needed to produce a unit of output in another, whereas the bridge coefficients indicate the product composition of the aggregate expenditures (GNP components).

In the third section of the paper, the general method of industrial forecasting in an I/O framework is illustrated in a simplified three-industry economy. The example shows how one could obtain the final demand by industry from a given set of GNP component forecasts. This example also illustrates how the I/O methodology enables us to take account of the chain of indirect effects (i.e., the ramifications across all industries of a change in the demand for the output of any particular industry).

In the forecasting procedure illustrated in the third section, it is assumed that the basic parameters of the model, namely the **A** and **H** matrices, will remain constant in the forecasting period. However, this assumption should be relaxed for making forecasts in a realistic environment where innumerable forces are at work causing variations in the basic parameters of the model.

There are a variety of methods for taking account of the changes in the **A** and **H** matrices directly or indirectly. An evaluation of a number of direct and indirect methods can be found in Reference 3. The direct methods that are usually employed to explain and predict the changes in each coefficient in the **A** and **H** matrices require the use of a large number of equations. This requirement presents several problems: First, the equations are to be estimated simultaneously with appropriate restrictions on their coefficients in order that the basic identities of the I/O frame-

work are not violated. The problems involved in estimating such equations for a large-scale I/O model are formidable as can be seen from the studies listed in References 4-7. Second, we do not have sufficient historical data to study and forecast each coefficient in the **A** and **H** matrices.

As the practical and theoretical difficulties of the direct methods seem to outweigh their advantages, we decided to adopt an indirect approach to take account of the changes in the **A** and **H** matrices in the IBM I/O model. Our method does not require a series of observations on each of the coefficients. The fourth section of the paper includes a discussion of how the changes in the **A** and **H** matrices are indirectly taken into account in the I/O model.

The I/O model is used to forecast current- and constant-dollar GPO for 86 industries and output for 51 manufacturing industries. Value added by industry, which is introduced in the first section, and Gross Product Originating, which is forecast by the I/O model, are conceptually very similar. Whereas we have a time series of GPO from the U.S. Department of Commerce, value added data by industry is available for only a selected number of years. As a result, we use GPO instead of value added for analysis and forecasting in the IBM I/O model. Also, the sources of basic data used for constructing these variables are not the same. The similarities and the differences between GPO and value added are discussed in the fourth section.

In the last section, the simulation performance of the I/O model is analyzed for the period 1955 to 1975.

The input-output transactions table

Every firm can be examined from two points of view: first, as a producer of the output it sells to other firms and to the final users of its product, and second, as a user of the inputs it buys from other firms and of the primary factors of production it purchases (labor, capital, space, etc.). If all business firms, households, and governments are grouped into industries, the same two-fold market structure still holds. Industries buy in one range of markets and sell in another. The I/O transactions table shows these dual market relationships among all industries in the economy.

An example of an I/O table that shows the flows of goods and services among different branches of an economic system during a particular period of time is given in Figure 1. In this simple example, there are only three producing industries - (1) Agriculture, (2) Manufacturing, and (3) Services. There are four final

Figure 1 An I/O transactions table (in dollars)

	Intermediate Demand (Buying Industries)				Final Demand (Final users)			Total Output
	Agr	Mfg	Svcs	Hshlds	Invest	Govt	Foreign	
	(1)	(2)	(3)	(C)	(I)	(G)	(EX)	
<i>Selling Industries</i>								
Agr(1)	—	20	45	30	—	3	2	100
Mfg(2)	30	10	20	10	40	60	30	200
Svcs(3)	—	80	—	60	—	5	5	150
				<u>100</u>	<u>40</u>	<u>68</u>	<u>37</u>	
<i>Value Added (VA)</i>								
	(1)	(2)	(3)					
Wages	20	50	20					
Profits	30	30	50					
Other	20	10	15					
Total VA	70	90	85					
Cost of materials + VA	100	200	150					
Total Value Added = \$70 + \$90 + \$85								
= \$245								
= GNP								
Total Final Demand = \$100 + \$40 + \$68 + \$37								
= \$245								
= GNP								

Figure 2 An I/O transactions table

Selling Industries	Intermediate Demand (Buying Industries)						Final Demand (Final Users)				Total Output
	1	2	...	j	...	N	C	I	G	EX	
1	x_{11}	x_{12}	...	x_{1j}	...	x_{1N}	f_{11}	f_{12}	f_{13}	f_{14}	X_1
2	x_{21}	x_{22}	...	x_{2j}	...	x_{2N}	f_{21}	f_{22}	f_{23}	f_{24}	X_2
...	...	...	...	...	...	...	...	...	...	...	...
i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{iN}	f_{i1}	f_{i2}	f_{i3}	f_{i4}	X_i
...	...	...	...	...	...	...	...	...	...	...	...
N	x_{N1}	x_{N2}	...	x_{Nj}	...	x_{NN}	f_{N1}	f_{N2}	f_{N3}	f_{N4}	X_N
Value Added	V_1	V_2	...	V_j	...	V_N					

users: Households, Investors, Government, and Foreign Countries. In the IBM I/O model, however, the number of industries (86) far exceeds the number of final users (16). (See the Appendix for a list of industries and final user categories of the IBM I/O model.) An I/O table for the general case of N industries is given in Figure 2.

Industry
sales

In Figure 1, the flows of goods and services among different industries are shown under the heading of "Intermediate Demand."

For example, the second row in the I/O table indicates that industry number 2 (Manufacturing) sells \$30 of output to industry 1 (Agriculture); it uses \$10 of its own product and sells \$20 to industry 3 (Services).

The transactions recorded under the heading "Final Demand" show the sales of the producing industries 1, 2, and 3 to various final users. From the second row it can be seen that industry 2 sells \$10 of its output to consumers, \$40 to investors, \$60 to governments, and \$30 to foreign countries. Thus for industry 2 we have:

$$\begin{aligned}\text{Intermediate demand} &= x_{21} + x_{22} + x_{23} \\ &= \$30 + \$10 + \$20 \\ &= \$60\end{aligned}\quad (1)$$

where

x_{21} represents sales by industry 2 to industry 1 (\$30).
 x_{22} represents use by industry 2 of its own output (\$10).
 x_{23} represents sales by industry 2 to industry 3 (\$20).

Total final demand for industry 2 is given by

$$\begin{aligned}F_2 &= f_{21} + f_{22} + f_{23} + f_{24} \\ &= \$10 + \$40 + \$60 + \$30 \\ &= \$140\end{aligned}\quad (2)$$

where

f_{21} is sales by industry 2 to consumers (\$10).
 f_{22} is sales by industry 2 to investors (\$40).
 f_{23} is sales by industry 2 to governments (\$60).
 f_{24} is sales by industry 2 to foreign countries (\$30).

By adding intermediate and final demands for any industry, shown in the row corresponding to that industry, we get a basic input-output identity. It is called the *row identity* and it gives the total value of sales for that industry.

$$\begin{aligned}\text{Total output} &= \text{intermediate demand} + \text{final demand} \\ \text{i.e., } X_2 &= x_{21} + x_{22} + x_{23} + F_2\end{aligned}\quad (3)$$

where X_2 is total output of industry 2 and the total final demand for its products is:

$$F_2 = f_{21} + f_{22} + f_{23} + f_{24}\quad (4)$$

If we have N industries and M final demand categories, the row identity for the i th industry will be:

$$X_i = \sum_{j=1}^N x_{ij} + F_i\quad (5)$$

where

$$F_i = \sum_{k=1}^M f_{ik} \quad (6)$$

**final
demand
identities**

The final demand identities show the product distribution of the GNP components. In Figure 1, for example, if we sum down the final user column (C), we will get the total amount of output sold by all the industries to the households for consumption; if we sum down the second final demand column (I), we will get total value of the machinery and other investment goods sold by all the industries to business firms. Similarly, the total amount of goods and services sold to governments by all the industries is obtained by summing down the column (G), and the total of net exports by all the industries is obtained by summing down the column (EX).

Denoting the totals for these four final demand categories as E_1, E_2, E_3 , and E_4 , we find that $E_1 = \$100$, $E_2 = \$40$, $E_3 = \$68$, and $E_4 = \$37$. Their sum totals to the GNP (\$245) of the economy. E_1, E_2, E_3 , and E_4 are often referred to as the GNP components. (The terms *GNP components*, *aggregate expenditures by category*, and *final demand totals by category* are used synonymously.)

**cost structure
of industries**

In Figure 1, the columns under "Intermediate Demand" show the materials purchased by different industries for use in their production processes. For example, reading down the second column, we have materials purchases by industry 2 (\$20 from industry 1, \$10 from industry 2, and \$80 from industry 3). The primary factor payments (wages, profits, interest, rent, etc.) by different industries constitute "Value Added". In Figure 1, from the second column under "Value Added," it can be seen that the wage payments of industry 2 are \$50, profits are \$30, and other factor payments are \$10. The total value added in industry 2 is \$90. If we sum the materials purchased by industry 2 from all other industries with the value added in industry 2, we will arrive at the total output of industry 2. Thus, the *column identity* for industry 2 is:

$$\begin{aligned} \text{Output} &= \text{Cost of materials} + \text{Value Added} \\ \text{i.e., } X_2 &= x_{12} + x_{22} + x_{32} + V_2 \end{aligned} \quad (7)$$

where V_2 is the value added in industry 2.

$$\begin{aligned} X_2 &= \$20 + \$10 + \$80 + \$90 \\ &= \$200 \end{aligned}$$

In the general case shown in Figure 2, the column identity is:

$$X_j = \sum_{i=1}^N x_{ij} + V_j \quad (8)$$

If we sum the value added in all the industries, we will get GNP, the total value of goods and services produced by the economy.

In the example given in Figure 1,

$$V_1 + V_2 + V_3 = \$70 + \$90 + \$85 = \$245 \quad (9)$$

In general,

$$V_1 + V_2 + \cdots + V_N = \text{GNP} \quad (10)$$

The above identity shows each industry's contribution to the total GNP. It can be seen from Figure 1 that, by summing the expenditure totals for the four final demand categories (Households, Investors, Governments, and Foreign Countries), we arrive at the same figure of \$245. Thus,

Total expenditure = Total factor payments

i.e.,

$$\sum_{k=1}^M E_k = \sum_{i=1}^N V_i = \text{GNP} \quad (11)$$

In the above identity, N refers to the number of industries, and M refers to the number of final demand categories.

The identity of Equation 11 holds even if we replace value added by GPO in each of the industries. As it was pointed out earlier in the paper, GPO by industry is very similar to value added by industry. It includes wages, salaries, profits, interest, rent, depreciation, etc. and excludes intermediate costs. When summed over all industries, it equals the economy's GNP.

I/O tables similar to those shown in Figures 1 and 2 are constructed by the U.S. Department of Commerce. An I/O table for the year 1958 covering 87 industries was first published by the U.S. Commerce Department during 1965. Since then, the Commerce Department has published a 1947 table consistent with the 1958 industry classification, a 1963 table with a 367-industry break, and a 1967 table with a 484-industry break.

The basic parameters of an input-output model

The basic parameters of an I/O model are the *input-output coefficients* and the *final demand coefficients*. These are usually estimated by simple ratios taken from a base year I/O table.

While the dollar value of flows among the industries is subject to considerable change from year to year, the coefficients obtained from the I/O table are much more stable. Input-output forecasting methods take advantage of this temporal stability.

The I/O coefficient, denoted by a_{ij} , shows the value of input needed from industry i to industry j to produce a dollar's worth of industry j 's output.

I/O coefficients

In general, the I/O coefficient is estimated as a ratio of the value of the input to the value of the total output of the industry at a point of time. That is,

$$\begin{aligned} a_{ij} &= x_{ij} \div X_j \\ &= i, j\text{th I/O coefficient} \end{aligned} \quad (12)$$

where

x_{ij} = Sale of i th industry's products to j th industry.

X_j = Total output of industry j .

From the I/O table presented in Figure 1, nine (3×3) I/O coefficients can be constructed as follows:

$$\begin{aligned} a_{11} &= 0.0 \div 100 \\ &= 0.0 \end{aligned}$$

$$\begin{aligned} a_{12} &= 20 \div 200 \\ &= 0.1 \end{aligned}$$

$$\begin{aligned} a_{13} &= 45 \div 150 \\ &= 0.3 \end{aligned}$$

...

$$\begin{aligned} a_{33} &= 0.0 \div 150 \\ &= 0.0 \end{aligned}$$

Arranging them in matrix form we get:

$$\mathbf{A} = \begin{bmatrix} 0.0 & 0.1 & 0.3 \\ 0.3 & 0.05 & 0.13 \\ 0.0 & 0.4 & 0.0 \end{bmatrix}$$

For example, a_{12} shows the value of agricultural products directly needed to produce a dollar's worth of a manufacturing industry's output. For the general case of N industries shown in Figure 2, we have

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{iN} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{N1} & a_{N2} & \cdots & a_{Nj} & \cdots & a_{NN} \end{bmatrix}$$

By substituting the relationship of the type given in Equation 12 into Equation 5, we can write the row identity for the i th industry as:

$$X_i = \sum_{j=1}^N a_{ij} X_j + F_i; i = 1, 2, \cdots, N \quad (13)$$

These equations can be written in matrix notation as:

$$\mathbf{X} = \mathbf{AX} + \mathbf{F} \quad (14)$$

where

$$\mathbf{X} = (X_1, X_2, \cdots, X_N)'$$

and

$$\mathbf{F} = (F_1, F_2, \dots, F_N)'$$

The second set of coefficients which are crucial in I/O forecasting are the *Final Demand Coefficients*. (These are also known as the *Bridge coefficients*.) The final demand coefficients show the product composition of various categories of aggregate expenditures (GNP components). The final demand coefficient denoted by h_{ik} shows the share of i th industry's products in the k th GNP component.

final demand
coefficients

$$h_{ik} = f_{ik} \div E_k \quad (15)$$

where

f_{ik} is the dollar value of i th industry's products going into the k th final demand category (i.e., k th GNP component).

E_k is the total value of the k th category of final demand.

In Figure 1, E_1 shows the total household expenditure for personal consumption and is equal to \$100. Out of this \$100, \$30 are spent on goods produced by industry 1 (agricultural products). Similarly, total spending for investment by all the economic units in the entire economy is E_2 and is equal to \$40. None of this \$40 is spent on goods produced by industry 1 because the agricultural sector does not produce any machinery or other investment goods. Therefore,

$$\begin{aligned} h_{11} &= 30 \div 100 \\ &= 0.3 \end{aligned}$$

$$\begin{aligned} h_{12} &= 0 \div 40 \\ &= 0.0 \end{aligned}$$

...

$$\begin{aligned} h_{34} &= 5 \div 37 \\ &= 0.135 \end{aligned}$$

Arranging them in a matrix form, we get:

$$\mathbf{H} = \begin{bmatrix} 0.3 & 0.0 & 0.044 & 0.054 \\ 0.1 & 1.0 & 0.882 & 0.811 \\ 0.6 & 0.0 & 0.074 & 0.135 \end{bmatrix}$$

In general,

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1M} \\ h_{21} & h_{22} & \dots & h_{2M} \\ \dots & \dots & \dots & \dots \\ h_{i1} & h_{i2} & \dots & h_{iM} \\ \dots & \dots & \dots & \dots \\ h_{N1} & h_{N2} & \dots & h_{NM} \end{bmatrix}$$

If the matrices $\mathbf{A}$ and $\mathbf{H}$ are reasonably stable over time, one may construct them for a base year and use them in making forecasts

of outputs by industry. However, technological and other factors cause changes in the **A** and **H** coefficients from year to year. While we do not have sufficient historical data to study the shifts in each of the $N^2 + N \times M$ coefficients, we must take these changes into consideration when forecasting.

In the IBM I/O model, we adopt an econometric approach to capture these shifts indirectly. Our method does not require direct observations on each of the coefficients for every year. The methodology of the IBM I/O model will be elaborated later.

Forecasting the outputs by industry

In order to understand the general methodology of input-output forecasting, it is worthwhile to explore how to make industrial projections if the **A** and **H** matrices remain constant through time. In a later section we will discuss how our forecasting strategy is modified to be applicable in a more realistic environment where the **A** and **H** matrices vary through time.

In reference to Figure 1, suppose we know the future values of the GNP components: specifically, that the projected values of total household expenditure for consumption, business investment, government expenditures, and net exports to foreign countries are \$200, \$60, \$120, and \$100, respectively. With this information, the **A** and **H** matrices are quite useful in estimating the industry-by-industry requirements of outputs implied in these GNP component projections.

The first step is to compute the industry distribution of these four categories of final demand aggregates. Applying the **H** matrix, the final demands for each industry's products are estimated as follows:

$$\mathbf{F} = \mathbf{H}\mathbf{E} \quad (16)$$

where

$$\begin{aligned} \mathbf{F} &= (F_1, F_2, F_3)' \\ \mathbf{E} &= (E_1, E_2, E_3, E_4)' \\ \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} &= \begin{bmatrix} 0.3 & 0.0 & 0.044 & 0.054 \\ 0.1 & 1.0 & 0.882 & 0.811 \\ 0.6 & 0.0 & 0.074 & 0.135 \end{bmatrix} \begin{bmatrix} \$200 \\ \$60 \\ \$120 \\ \$100 \end{bmatrix} \end{aligned} \quad (17)$$

For example, in arriving at the final demand for industry 1, the calculations in Equation 17 imply the following assumptions:

- Of the \$200 of total expenditure by households for consumption, 30 percent will be purchased from industry 1 (i.e., agricultural products).

- Of the \$60 of investment expenditures by the entire economy, nothing will be spent on agricultural products.
- Similarly, 4.4 percent of government expenditures will be on the products produced by agriculture.
- 5.4 percent of exports will consist of agricultural products.

Thus, the total final demand for the products of industry 1 will be:

$$\begin{aligned} F_1 &= (0.3 \times \$200) + (0.0 \times \$60) + (0.044 \times \$120) \\ &\quad + (0.054 \times \$100) \\ &= \$70.7 \end{aligned}$$

Similarly, the final demands for industries 2 and 3 can be calculated as:

$$\begin{aligned} F_2 &= (0.1 \times \$200) + (1.0 \times \$60) + (0.882 \times \$120) \\ &\quad + (0.811 \times \$100) \\ &= \$267.0 \end{aligned}$$

and

$$\begin{aligned} F_3 &= (0.6 \times \$200) + (0.0 \times \$60) + (0.074 \times \$120) \\ &\quad + (0.135 \times \$100) \\ &= \$142.3 \end{aligned}$$

The next task is to estimate the outputs needed directly and indirectly from each of the three industries to support the final demands we just computed.

Let the total (direct and indirect) output requirements from industries 1, 2, and 3 be X_1 , X_2 , and X_3 respectively. In order to produce X_1 , X_2 , and X_3 dollars of output by industries 1, 2, and 3 the intermediate demand requirements will be:

$$AX = \begin{bmatrix} 0.0 & 0.1 & 0.3 \\ 0.3 & 0.05 & 0.13 \\ 0.0 & 0.4 & 0.0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad (18)$$

For example, the intermediate demand for the products of industry 1 consists of flows to itself, flows to industry 2, and flows to industry 3. According to the equations of (18):

- To produce X_1 dollars of output, industry 1 needs $a_{11} \times X_1 (= 0.0 \times X_1)$ dollars' worth of its own products.
- To produce X_2 dollars of output, industry 2 requires $a_{12} \times X_2 (= 0.1 \times X_2)$ dollars' worth of products from industry 1.
- To produce X_3 dollars of output, industry 3 needs $a_{13} \times X_3 (= 0.3 \times X_3)$ dollars' worth of output from industry 1.

Thus the output required from industry 1 for intermediate use by all the industries is:

$$a_{11} \times X_1 + a_{12} \times X_2 + a_{13} \times X_3$$

Since we know from the A matrix that $a_{11} = 0$, $a_{12} = 0.1$, and $a_{13} = 0.3$, the intermediate demand facing industry 1 is:

$$0.0 \times X_1 + 0.1 \times X_2 + 0.3 \times X_3$$

Similarly, the equations of (18) state that the intermediate demand for the products of industries 2 and 3 are:

$$0.3 \times X_1 + 0.05 \times X_2 + 0.13 \times X_3$$

and

$$0.0 \times X_1 + 0.4 \times X_2 + 0.0 \times X_3, \text{ respectively.}$$

The total output requirements from each industry are obtained by summing intermediate and final demands. Therefore,

$$\mathbf{X} = \mathbf{AX} + \mathbf{F} \quad (19)$$

where $\mathbf{AX}$ is a vector of intermediate demands and $\mathbf{F}$ is a vector of final demands, i.e.,

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 0.0 & 0.1 & 0.3 \\ 0.3 & 0.05 & 0.13 \\ 0.0 & 0.4 & 0.0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} \quad (20)$$

The above equations imply that the total output of industries 1, 2, and 3 are:

$$\begin{aligned} X_1 &= (0.0 \times X_1) + (0.1 \times X_2) + (0.3 \times X_3) + F_1 \\ X_2 &= (0.3 \times X_1) + (0.05 \times X_2) + (0.13 \times X_3) + F_2 \\ X_3 &= (0.0 \times X_1) + (0.4 \times X_2) + (0.0 \times X_3) + F_3 \end{aligned}$$

Rewriting these equations, we get

$$(\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{F} \quad (21)$$

where $\mathbf{I}$ is the identity matrix,

$$\mathbf{I} = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

Expanding Equation 21,

$$\begin{bmatrix} (1.0 - 0.0) & -0.1 & -0.3 \\ -0.3 & (1.0 - 0.05) & -0.13 \\ -0.0 & -0.4 & (1.0 - 0.0) \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} \quad (22)$$

$$\begin{aligned} (1 - 0.0) \times X_1 - 0.1 \times X_2 - 0.3 \times X_3 &= F_1 \\ -0.3 \times X_1 + (1 - 0.05) \times X_2 - 0.13 \times X_3 &= F_2 \\ -0.0 \times X_1 - 0.4 \times X_2 + (1.0 - 0.0) \times X_3 &= F_3 \end{aligned} \quad (23)$$

We know from Equation 17 that $F_1 = \$70.7$, $F_2 = \$267.0$, and $F_3 = \$142.3$. By substituting them in the equations of (22) or (23) and solving for the three unknowns, X_1 , X_2 , and X_3 , we get:

$$X_1 = \$198.1, X_2 = \$385.2, \text{ and } X_3 = \$296.4$$

In matrix notation the above solution is represented as:

$$\mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{F} \quad (24)$$

The general method of I/O forecasting (demonstrated above) consists of two main steps, which can be generalized for the case of N industries and M GNP components as follows. (The IBM I/O model has more than the two basic steps described here. A full description of the I/O model is presented in the next section.)

a. Calculation of the final demands by industry:

$$F_i = \sum_{k=1}^M h_{ik} E_k; i = 1, 2, \dots, N \quad (25)$$

where E_k is the total value of k th final expenditure category (i.e., k th GNP component).

In matrix notation the equations of (25) become:

$$\mathbf{F} = \mathbf{H}\mathbf{E} \quad (26)$$

where

$$\mathbf{F} = (F_1, F_2, \dots, F_N)'$$

and

$$\mathbf{E} = (E_1, E_2, \dots, E_M)'$$

b. Solving for the output requirements by industry:

To calculate total requirements, the equations of (13) are used, i.e.,

$$X_i = \sum_{j=1}^N a_{ij} X_j + F_i; i = 1, 2, \dots, N \quad (27)$$

These equations can be written in matrix notation as:

$$\mathbf{X} = \mathbf{A}\mathbf{X} + \mathbf{F} \quad (28)$$

where

$$\mathbf{X} = (X_1, X_2, \dots, X_N)'$$

and

$$\mathbf{F} = (F_1, F_2, \dots, F_N)'$$

The total output requirements by industry are obtained by solving Equation 28 as follows:

$$\mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{F} \quad (29)$$

$(\mathbf{I} - \mathbf{A})^{-1}$ is known as the total requirements matrix. The i, j th element of this matrix represents the amount of output of industry i required directly and indirectly to satisfy one dollar's worth of final demand for industry j . In the example of Figure 1,

$$(\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} 1.079 & 0.265 & 0.359 \\ 0.361 & 1.204 & 0.269 \\ 0.144 & 0.482 & 1.108 \end{bmatrix}$$

For example, to produce agricultural products we need fertilizers, gasoline, etc., from the manufacturing sector. Therefore, if final

demand for agricultural products increases by \$100, initially it will generate an indirect demand of \$30 for manufacturing output (see the A matrix given earlier in the discussion on I/O coefficients). To produce this \$30 of manufacturing output, an additional \$3 of output from agriculture, \$1.5 of output from manufacturing, and \$12 of output from services are needed. Thus a chain of indirect effects is generated. Taking all these indirect effects into consideration, we will need \$107.9 of agricultural products, \$36.1 of manufacturing output and \$14.4 of services, to support the \$100 of final demand for agricultural products. These figures can be obtained from the $(I - A)^{-1}$ matrix given above. In Equation 29, the total requirements are implicitly calculated as the sum of direct and all indirect effects and can be seen from the power series expansion of the $(I - A)^{-1}$ matrix:

$$\begin{aligned} X &= (I - A)^{-1}F \\ &= F + AF + A^2F + A^3F + \cdots + A^nF + \cdots \end{aligned} \quad (30)$$

The IBM input-output model

In the previous section, a methodology was developed for projecting the output requirements by industry, given future values of the GNP components. It was shown that such a forecasting method relies heavily on the A and H matrices. The simple example shown in the previous section involved only three industries and used only the base year A and H matrices to arrive at its output forecasts. However, base year A and H matrices can provide only preliminary forecasts (or first approximations) of realistic outputs by industry. These predictions must be refined substantially in order to account for changes in the A and H matrices. The IBM I/O model makes provisions for such refinements for each of its 86 industries.

output
forecasts
by industry

In the IBM I/O model, preliminary forecasts of outputs by industry are obtained as follows:

$$\hat{X}_{T+\tau} = (I - A_0)^{-1}F_{T+\tau} \quad (31)$$

and

$$F_{T+\tau} = H_0 E_{T+\tau} \quad (32)$$

These equations are the same as those of Equations 24 and 26 except that the time subscripts are attached to the variables and to the A and H coefficients. That is, A_0 is the I/O coefficient matrix for the base period (i.e., 1967) and H_0 is the base year final demand coefficient matrix. T is the current period and $T + \tau$ is τ years ahead into the future. $E_{T+\tau}$ is the vector of GNP components for the forecast period $T + \tau$:

$$E_{T+\tau} = (E_1, E_2, \cdots, E_M)'_{T+\tau}$$

$F_{T+\tau}$ is the final demand vector for the period $T + \tau$:

$$\mathbf{F}_{T+\tau} = (F_1, F_2, \dots, F_N)'_{T+\tau}$$

and $\hat{\mathbf{X}}_{T+\tau}$ is the vector of preliminary forecasts of outputs by industry for the period $T + \tau$:

$$\hat{\mathbf{X}}_{T+\tau} = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_N)'_{T+\tau}$$

In order to arrive at a methodology for correcting the preliminary forecasts, we compared them with the actuals in the historical period. Given the base year (1967) $\mathbf{A}$ and $\mathbf{H}$ matrices and the historical time series of the GNP components, a set of synthetic outputs (preliminary output estimates) were obtained for the period 1950–1975 as follows:

$$\hat{\mathbf{X}}_t = (\mathbf{I} - \mathbf{A}_0)^{-1} \mathbf{H}_0 \mathbf{E}_t \quad (33)$$

t refers to the years 1950 through 1975 and 0 refers to the base year, 1967. $\hat{\mathbf{X}}_t$ is the vector of synthetic outputs by industry for the period t .

$\hat{X}_{it}$ is the synthetic output for the i th industry during the period t , and X_{it} is the actual output for the i th industry in the same period. The discrepancy $X_{it} - \hat{X}_{it}$ shows the changing importance of the i th industry in the economy. If $X_{it} - \hat{X}_{it}$ is greater than zero, it means that the importance of the i th industry is greater in period t than in the base period 1967. For the industries with declining importance, $X_{it} - \hat{X}_{it}$ shows a downward trend. Such had been the case for the coal mining industry until the early seventies. Since then, with the onset of crude oil and natural gas shortages, the discrepancy has started to move up. Similarly, the growing industries like plastics show an upward trend.

During the period 1950–75, the discrepancies between the actual and synthetic outputs showed long-term trends for some industries and cyclical movements for others, and some of the industries exhibited both types of movements. In addition, the discrepancies were found to be correlated with variables such as time trend, GNP, unemployment rate, etc. These correlations are hardly surprising, because $X_{it} - \hat{X}_{it}$ is mainly affected by movements in the $\mathbf{A}$ and $\mathbf{H}$ matrices, caused by such factors as technological advance, changes in tastes, material shortages, rates of capacity utilization, and the general economic conditions.

Technological changes and changes in tastes and quality tend to occur in a gradual way through time. Hence, a time trend serves as a good explanatory variable for capturing these gradual changes. Since the efficiency of raw material use tends to vary with supply as well as demand pressures, such variables as rate of capacity utilization are also needed to explain variations in the discrepancies.

The elements of the $\mathbf{H}$ matrix are also sensitive to both short-term and long-term factors. Thus, economic indicators such as GNP, unemployment rate, changes in inventories, etc. and a time trend are used to explain the discrepancies between the preliminary forecasts and the actuals. Equations of the following form can be used to explain and forecast these discrepancies:

$$X_{it} - \hat{X}_{it} = \alpha_i + \delta_i Z_t + \gamma_i t \quad (34)$$

where Z_t represents the economic variables such as GNP or unemployment rate and t represents the time trend.

The above form, however, is not used for many industries in the model because the time series of outputs and the benchmark outputs given in the 1967 I/O table differ somewhat in their coverages. If we assume that the outputs generated by using the base year $\mathbf{A}$ and $\mathbf{H}$ matrices are to be scaled up or down in order to make them consistent with the time series data, we can use the following alternative form of Equations 34:

$$X_{it} = \alpha_i + \beta_i \hat{X}_{it} + \delta_i Z_t + \gamma_i t \quad (35)$$

The β_i coefficient turned out to be close to unity in those industries where the differences in the time series data and the benchmark data (1967 I/O table) are negligible.

Also, the economic variable (Z_t) is not necessarily the same for all the industries. For some industries GNP gives better explanations; for others the unemployment rate is better. After a great deal of testing, the appropriate explanatory variables were selected for each of the 86 industries.

The final output forecast for the i th industry is obtained from:

$$\hat{X}_{iT+\tau} = \alpha_i + \beta_i \hat{X}_{iT+\tau} + \delta_i Z_{T+\tau} + \gamma_i (T + \tau)$$

**GPO forecasts
by industry**

The U.S. Department of Commerce constructs a time series of *Gross Product Originating* (GPO) for a number of manufacturing and nonmanufacturing industries. GPO is conceptually very close to "value added" by industry, which is described in Figure 1. Like value added, it includes primary factor payments and excludes intermediate raw material costs. It includes such items as employee compensation, profits, depreciation, etc. Also, the sum of GPO in all the industries is equal to total GNP in the economy.

However, the source of the basic data used in the construction of value added is different from that used for GPO. Value added as given in the I/O tables published by the U.S. Department of Commerce⁸ is obtained from the economic census, and it is available for only a selected number of years during the historical period. Also, value added in the I/O tables is often derived as a residual, i.e., as the difference between total output and cost of materials and services.

GPO by industry is available for all the years between 1950 and 1975. The Commerce Department constructs the GPO series by adding each industry's share of capital consumption and other adjustment items to income originating in that industry. One major source for the wage and salary component of income originating by industry is the state unemployment agencies. The profit component for many industries is obtained from Internal Revenue Service data. The details of other sources and methods used for constructing the GPO series is available in Reference 9.

The IBM I/O model predicts GPO for each of the 86 industries. In the model, GPO forecasts are derived from the output forecasts using the GPO/output ratios.

From historical experience, we have found that the GPO/output ratios by industry are sensitive to the level of general economic activity and specifically to variables such as corporate profits. For some industries, a slight trend is also observed. To forecast the GPO/output ratios, equations of the following type are used:

$$\frac{\text{GPO}_{it}}{X_{it}} = \Psi_i(t, \text{CYCLE}_t) \quad (37)$$

where t is a time trend variable and CYCLE is a cyclical variable such as corporate profits before taxes. The GPO forecasts for the i th industry are obtained as:

$$\text{GPO}_{iT+\tau} = X_{iT+\tau} \times \Psi_i(T + \tau, \text{CYCLE}_{T+\tau}) \quad (38)$$

Time series of outputs are not readily available for most of the nonmanufacturing industries, so for nonmanufacturing industries we forecast GPO directly without going through the output forecasts.

The IBM I/O model is built in value terms, not in physical terms. That is, the **A** and **H** coefficients used in this model are constructed from dollar values of flows of goods and services among the industries and between the industries and the final users. The forecasting is also carried out in value terms. Starting with a set of projected values of current dollar GNP components, forecasts of output and GPO are obtained in current prices. In order to arrive at the constant dollar GPO forecasts, we need projections of the deflators. For each industry, predictions of the GPO deflators are obtained using equations of the type:

$$P_{it} = \Phi_i(P_t, \text{ULC}_t, C_t) \quad (39)$$

where P_{it} stands for the GPO deflator of the i th industry and P_t stands for an aggregate price variable such as the GNP deflator, Wholesale Price Index, or some other aggregate price index. ULC_t is an indicator of the unit labor costs in the economy and C_t is a cyclical variable such as the unemployment rate.

**GPO deflators
by industry**

Table 1 Simulation errors¹

<i>Year</i>	<i>GPO (curr²)</i>	<i>GPO (const³)</i>	<i>Output (curr²)</i>
1955	4.632	4.967	5.091
1956	4.098	4.872	4.777
1957	3.330	5.140	4.104
1958	3.954	5.420	4.602
1959	3.118	3.672	3.345
1960	3.426	4.668	3.452
1961	3.573	4.572	3.676
1962	3.518	3.801	3.423
1963	3.988	3.722	3.500
1964	4.383	4.193	3.864
1965	4.176	4.650	4.598
1966	3.965	4.886	4.485
1967	3.704	4.493	4.433
1968	4.124	4.703	4.727
1969	3.062	3.775	3.899
1970	3.153	3.490	4.474
1971	3.503	3.240	4.574
1972	3.150	4.020	4.231
1973	3.927	4.166	3.817
1974	4.950	4.603	4.644
1975	3.747	4.759	4.728

1. Weighted averages of absolute percentage errors by industry.

2. Curr: Measured in current prices.

3. Const: Measured in constant prices.

The forecast of the GPO deflator for the i th industry is obtained from:

$$P_{i\ T+\tau} = \Phi_i(P_{T+\tau}, \text{ULC}_{T+\tau}, C_{T+\tau}) \quad (40)$$

Constant dollar GPO forecasts are obtained by applying the projected deflators to the current dollar GPO forecasts for each of the 86 industries.

Simulation performance by the IBM input-output model

All economic forecasting models, no matter how well they are structured, are subject to some degree of error due to random and other unaccounted factors. In order to get an idea of the extent of error that the IBM I/O model is expected to make, we have simulated current dollar outputs and both current- and constant-dollar GPO by industry for the period 1955–1975.

In general, the simulation error for any variable is computed as:

$$e_{it} = A_{it} - P_{it}$$

where

P_{it} is the simulated value of the variable for the i th industry in period t .

A_{it} is the actual value of the variable for the i th industry in period t .

e_{it} is the simulation error in the i th industry in period t .

The Absolute Percentage Error (APE) for the i th industry at time t is defined as

$$APE_{it} = \frac{|e_{it}|}{A_{it}} \times 100 \quad (41)$$

The Average Absolute Percentage Error (AAPE) for all the industries in period t is:

$$AAPE_t = \frac{\sum_{i=1}^N w_{it} APE_{it}}{\sum_{i=1}^N w_{it}} \quad (42)$$

where w_{it} is the weight of the i th industry in period t . The weight of the i th industry is its output or GPO, as the case may be. The AAPES in the output simulations and in the current- and constant-dollar GPO simulations are shown in Table 1.

In the first section, it was shown that the sum of value added (or GPO) for all industries is equal to the GNP of the economy. That is,

$$\sum_{i=1}^N V_{it} = GNP_t \quad (43)$$

Table 2 shows the absolute difference between $\sum \hat{V}_{it}$ and GNP_t , expressed as a percentage of GNP_t , for the period 1956–78. $\hat{V}_{it}$ is the GPO simulation for the i th industry in period t . This table shows how well the identity in Equation 43 holds in current and constant dollars during the simulation period as well as in the forecast period.

ACKNOWLEDGMENT

I would like to thank Dr. Joel Brest of the Economic Research Department, IBM Corporate Headquarters, for many helpful comments on this paper.

CITED REFERENCES

1. *IBM Annual Econometric Model*, Economic Research Department, IBM Corporation, Armonk, NY (July 1976).
2. "Input-output tables and analysis," *Studies in Methods*, Series F, No. 14, United Nations, New York, NY (1973).
3. K. S. Sarma, *Comparative Performance of Input-Output Models with Alternative Production Functions*, Unpublished Doctoral Dissertation, Department of Economics, University of Pennsylvania, Philadelphia, PA (1972).

Table 2 Absolute percentage errors in GNP simulations*

Year	Curr	Const
1956	0.280	2.627
1957	0.394	3.763
1958	1.694	4.127
1959	1.077	2.667
1960	1.757	3.777
1961	1.775	2.905
1962	1.190	1.245
1963	1.007	4.231
1964	0.677	0.546
1965	0.314	1.403
1966	0.120	1.652
1967	0.209	0.720
1968	0.061	0.689
1969	0.407	0.953
1970	0.656	1.703
1971	0.365	0.106
1972	0.061	1.334
1973	0.545	1.887
1974	2.025	1.825
1975	1.177	1.088
1976	0.253	0.629
1977	0.169	1.321
1978	0.300	1.496

*The absolute difference between $\sum \hat{V}_{it}$ and GNP_t is expressed as a percentage of GNP_t .

4. K. Arrow and M. Hoffenberg, *A Time Series Analysis of Inter-industry Demands*, North Holland Publishing Company, Amsterdam, Netherlands (1959).
5. D. W. Jorgenson and E. A. Hudson, "U.S. energy policy and economic growth, 1975-2000," *The Bell Journal of Economics and Management Science*, 5, No. 2 (Autumn 1974).
6. R. S. Preston, "The Wharton long-term model: Input-output within the context of a macro forecasting model," *International Economic Review*, 16, No. 1 (February 1975).
7. K. S. Sarma, "Comparative study of alternative production functions in an input-output framework," Paper presented to the Third World Congress of the Econometric Society, Toronto, Canada (August 1975).
8. "The input-output structure of the U.S. economy, 1967," *Survey of Current Business*, U.S. Department of Commerce, The Interindustry Economics Division, Washington, DC (February 1974).
9. *GNP by Major Industries: Concepts and Methods*, U.S. Department of Commerce, Office of Business Economics, Washington, DC (April 1966).

GENERAL REFERENCES

1. L. R. Klein, "On the interpretation of Professor Leontief's system," *Review of Economic Studies*: XX (1952-53).
2. W. W. Leontief, *The Structure of American Economy, 1919-1939*, Oxford University Press, New York, NY (1951).

Appendix: List of industries and GNP components of the IBM I/O model

Table A1 SIC content of the IBM I/O model

<i>IBM sector</i>	<i>SIC (1967 edition)</i>
1 Agriculture, livestock	01
2 Forestry, fishery & agric. services	07 (excl. 0722, pt. 0729), 08, 09
3 Metal mining	10
4 Coal mining	11, 12
5 Crude petroleum & natural gas	1311, 1321
6 Stone, clay & mineral mining	14
7 Construction	15, 16, 17, pt. 138, pt. 6561
8 Food manufacturing	20
9 Tobacco	21
10 Fabrics	22 (excl. 225, 227, 229)
11 Misc. textile goods	227, 229
12 Apparel	225, 23 (excl. 239), 39996
13 Misc. fabricated textiles	239
14 Lumber & wood products	24 (excl. 244)
15 Wooden containers	244
16 Household furniture	251
17 Other furniture & fixtures	252-9
18 Paper & allied products	26 (excl. 265)
19 Paperboard containers & boxes	265

<i>IBM sector</i>	<i>SIC (1967 edition)</i>
20 Printing & publishing	27
21 Chemicals & products	281 (excl. 28195), 286-9
22 Plastics & synthetics	282
23 Drugs, cleaning, & toilet preparations	283, 284
24 Paints & allied products	2851
25 Petroleum refining & related industries	29
26 Rubber & miscellaneous plastics	30
27 Leather tanning & industrial leather	311, 312
28 Footwear & other leather products	313-9
29 Glass & glass products	3211, 3221, 3229, 3231
30 Stone & clay products	32 (excl. 3211, 3221, 3229, 3231)
31 Primary iron & steel mfg.	331, 332, 3391, 3399
32 Primary nonferrous metals mfg.	333-6, 3392, 28195
33 Metal containers	341, 3491
34 Heating, plumbing & fabricated metals	343, 344
35 Screw machine products & metal stamps	345, 346
36 Other fabricated metal products	342, 347-9 (excl. 3491)
37 Engines & turbines	351
38 Farm machinery	352
39 Construction, mining, oil field equip.	3531-3
40 Materials handling M&E	3534-7
41 Metal working M&E	354
42 Special industry M&E	355
43 General industrial M&E	356
44 Machine shop products	359
45 Office computing & acct. machinery	357
46 Service industry machines	358
47 Electric transmission & distribution equip.	361-2
48 Household appliances	363
49 Electric lighting & wiring equip.	364
50 Radio, television & comm. equip.	365-6
51 Electronic components & accesories	367
52 Misc. electrical M&E supplies	369
53 Motor vehicles & equipment	371
54 Aircraft & parts	19, 372
55 Other transportation equip.	373-9
56 Scientific instruments	381, 382, 384, 387
57 Optical & photographic equip.	383, 385, 386
58 Misc. mfg.	39 (excl. 39996)
59 Railroads	40, 474
60 Local highway & transit	41
61 Motor freight & warehousing	42, 473
62 Water transportation	44
63 Air transportation	45
64 Pipeline transportation	46
65 Transportation services	47 (excl. 473, 474)

<i>IBM sector</i>	<i>SIC (1967 edition)</i>
66 Communications	48 (excl. 483)
67 Radio & TV broadcasting	483
68 Utilities	49
69 Wholesale & retail trade	50-59, 7396, pt. 8099
70 Banking	60
71 Credit agencies	61, 67
72 Security & commodity brokers	62
73 Insurance	63, 64
74 Real estate & rental	65 (excl. pt. 6561), 66
75 Hotels & personal services	70, 72, 76 (excl. 7692, 7694, pt. 7699)
76 Business services	73 (excl. 7396), 7692, 7694, pt. 7699, 81, 89 (excl. 8921)
77 Auto repair & services	75
78 Amusements	78, 79
79 Medical & educational services	0722, 80 (excl. pt. 8099), 82
80 Nonprofit organizations	84, 86, 8921
81 Fed. govt. enterprises	
82 S & L govt. enterprises	
83 Imports	
84 Government industry	
85 Household industry	
86 Rest of the world industry	

Table A2 GNP components in the IBM I/O model

PERSONAL CONSUMPTION EXPENDITURE
Nondurable goods
Services
Automobiles and parts
Durable goods less autos & parts
GROSS PRIVATE DOMESTIC INVESTMENT
Residential structures
Nonresidential structures
Producers' durable equipment
Change in business inventories
EXPORTS OF GOODS AND SERVICES
IMPORTS OF GOODS AND SERVICES
GOVERNMENT EXPENDITURE
Federal structures
Federal purchases
Federal compensation
State and local structures
State and local purchases
State and local compensation

Reprint Order No. G321-5060.