Rensselaer
Working Papers in Economics

Mathematical Models in Input-Output Economics

Faye Duchin
Rensselaer Polytechnic Institute

Albert E. Steenge
University of Twente

Number 0703

April 2007

_____

Mathematical Models in Input-Output Economics

by

Faye Duchin
Department of Economics, Rensselaer Polytechnic Institute, Troy, New York, USA

Albert E. Steenge
School of Management and Governance, University of Twente, Enschede, The Netherlands

Abstract

This paper describes the mathematical basis for input-output economics, the major types of models, and the underlying economic theory. The features of these models that make them especially well suited for understanding the connections between the economy and the environment are emphasized throughout. These include the dual physical and price representations and the representation of resource inputs as factors of production, whether they are priced or not.

The basic static physical and price models are described, along with their major properties and associated databases.   The most important approaches to analysis involve multipliers, decomposition, and scenario analysis. Going beyond the basic static framework requires the progressive closure of the model by making exogenous variables endogenous while maintaining simplicity, transparency, and the distinctive feature of an input-output model: the simultaneous determination of solutions at the sectoral level and the economy-wide level. Closures for household activities and for investment are described by way of example.

The major extensions of the basic model accommodate the representation of pollutant emissions and policies for constraining them, dynamic models, and multi-regional models, the latter including a new version of a world model that solves for bilateral trade flows and region-specific prices based on comparative advantage with factor constraints. The concluding section describes the challenges currently being addressed within the field. An annotated bibliography provides references for further reading and includes both classic articles and a representation of recent research.

JEL Codes: C67, Q01, F1, O1, O3

Mathematical Models in Input-Output Economics

1. The Basic Static Input-Output Model

1.1. Introduction

Economists are concerned with promoting innovation, and achieving efficiency by reducing production costs, in order to maximize the prospects for growth, profits, and increased consumption. They portray the economy in terms of a circular flow of income between producers and consumers: producers pay incomes to workers, and workers use their income to buy goods and services. Goods are assumed to move around the circle in the opposite direction from the money flows.

While stylized, this image accurately conveys the duality between the systems of physical flows and of money values that constitute an economy. However, while the money flows indeed stay within the economic system, producing the physical flows requires inputs of resources from the environment and discharges wastes into the environment. Typically the role of the environment is ignored by economists: resources are treated as "free gifts of nature," and wastes are called "externalities," meaning that they are external to the economic system.

Input-output economics shares with other conceptual approaches to economics a concern for achieving efficiency in the use of resources to produce goods. However, input-output economics also accommodates other objectives, not only conceptually but also operationally: it differs from other schools of thought in not *imposing* growth and efficiency objectives on its mathematical models. Both resource inputs and wastes generated are standard variables in an input-output analysis that are explicitly represented in equations, whether they have prices or not. These features, while not always fully exploited, make input-output economics especially suitable for environmental analysis.

In the 1930s Wassily Leontief published a pair of articles that laid the groundwork for input-output economics. The first article described the design and construction of what he called a Tableau Economique of the United States for the year 1919, the direct precursor of today's input-output table with its distinctive focus on the inter-industry transactions until then essentially ignored by economists. The second article was in two parts. First came the theoretical framework, describing the interdependence of the different parts of an economy by a set of mathematical equations intended for manipulating this new kind of table, followed by an empirical implementation of the mathematical model. While the table had been compiled in terms of 41 sectors, it was for practical reasons aggregated to only 10 and, even so, the results had to be approximated because it was not then feasible to compute the inverse of a 10 x 10 matrix. The close integration of model and database remains a defining characteristic of input-output economics, even as both have been extended into new domains.

The order of the publications emphasized the fundamental role accorded to the table, comprised as it still is today of a square inter-industry portion, recording the flows of deliveries from each industry to every other one, supplemented by additional rows representing other inputs, namely labor and capital or, on occasion, resources such as water and land. The latter are designated factors of production, or factor inputs, to indicate that unlike other inputs they are not produced in any industry (or, in the case of capital goods, their capacity cannot be expanded in a single production period). The table is completed by additional columns that represent deliveries from industries to households and other final users; these are called final deliveries or final demand. The resulting table is rectangular, since the number of final demand categories is in general not equal to the number of factors of production. This rectangular table records all transactions taking place in the economy in a specific period of time. The capacity to absorb a substantial level of detail, and the conceptual simplicity and transparency of the framework, make the input-output table and the models that manipulate it well suited to evaluating strategies for sustainable development.

The reference in Leontief's article is to the 18th century Tableau Economique of François Quesnay, which depicted the flows of income among landlords, manufacturers (artisans), farmers and other agricultural workers, and merchants derived from the sale of agricultural goods and fabricated products. By contrast, a contemporary input-output flow table distinguishes dozens if not hundreds of sectors producing goods and services in a modern economy.

Input-output tables are compiled in many countries by official statistical offices, specialized official or semi-official institutes such as national banks or universities, private companies, or individual researchers. They use several methods to build input-output tables, all of which share characteristics associated with the System of National Accounts. These guidelines guarantee internal consistency of the tables, consistency with widely used national aggregates such as gross national product, and comparability among tables representing different economies. International organizations such as the UN, OECD (previously OEEC) or Eurostat have played an important role in issuing guidelines for the national accounts, including for input-output tables.

The square portion of the input-output table has n rows and n columns, and the figure in the ith row and j^{th} column represents the amount of product from industry i delivered to industry j in a particular calendar year. The result of dividing that quantity by the total output of industry j is a coefficient measuring input per unit of output. In this way the n x n portion of the flow table is converted to an n x n matrix of coefficients, of which the entries in the jth column include (when supplemented by the jth column of factor inputs per unit of output) all inputs needed to produce one unit of output of industry j. This column of coefficients is said to represent the average technology in use in industry j. For simplicity it is assumed that every industry, or sector, is associated with a single characteristic output produced using a single average technology.

Call the (n x n) matrix of interindustry coefficients A, the (n x 1) vector of outputs x, and the likewise (n x 1) vector of final deliveries y, while F is the k x n matrix of factor inputs

per unit of output (one row for each of k factors) and total factor use is the vector f. Then the basic static input-output model states that:

$(I - A) x = y$, or

$x = (I - A)^{-1}y$ and   (1)

$f = Fx$ (2)

where the inverse matrix $(I - A)^{-1}$ has been called the Leontief inverse. It is also known as the multiplier matrix or matrix of multipliers. (Because the economy needs to produce a larger amount of a specific good than the amount of final demand for that good, final demand y needs to be 'multiplied' to obtain x.)

Equations (1) and (2) are called a quantity input-output model, and the corresponding parameters (the coefficients in the A and F matrices) are ratios of physical units such as tons of plastic per computer (dollars' worth being one special case of a quantity). If y is given, the solution vector x represents the quantities of sectoral outputs. Equations (1) and (2) comprise the basic static input-output model. The following sections of this paper describe extensions to that model and the kinds of questions the extended models have been designed to answer.

1.2. Reasons for Popularity

Over the past several decades the basic static input-output model has been used extensively in empirical economic analysis. In the last decade, the special appeal of this approach for environmental analysis has become apparent for reasons that will be discussed below. In the typical case, a rectangular input-output flow table compiled by a statistical office for some past year is available as the starting point for deriving the coefficient matrices, A and F, and Equations (1) and (2) are used to compute the impact on outputs (x) or on employment (a component of f), of alternative hypothetical assumptions about changes in input coefficients (A and F) or in final deliveries (y). When there is no table, the matrices can be estimated directly from technological information or by starting from matrices for a similar economy. The fact that input-output tables are widely available makes the model easy to implement and, in turn, the popularity of the model has encouraged the production of tables that are increasingly detailed and frequent as well as more comprehensive in their coverage, notably including environmental data classified in categories that are compatible with input-output accounts.

Two important assumptions underlie the model given by Equations (1) and (2). First, output is a linear function of final demand. Second, in the absence of exogenous assumptions to the contrary, the input coefficients of matrices A and F remain constant for variations in y. As a consequence of these assumptions, if final demand for all commodities increased by 10%, total required outputs from each sector would also increase by this same percentage. A rationale for the latter, so-called fixed coefficients

assumption is that the columns of A (and the corresponding columns of F) represent the most efficient technologies (or production functions) available to produce each good, and they are assumed to remain the optimal ones even if there are variations in the composition of final demand. A more realistic rationale is that, while these technologies are not necessarily optimal, they are the ones effectively in place and cannot be quickly changed given the existing stock of fixed capital. In either case, the time period during which real-world technologies will in fact remain the same is limited because new technologies may become available over time and new fixed capital can be put in place. However, for a certain period of time the coefficients can be expected to remain more or less unchanged, and the model can be used to compute the required change in x even if y changes. (We shall come back to the subjects of linearity and fixed coefficients, starting with Sections 1.2 and 1.5).

The power of this simple model resides in the fact that, while an entry of A quantifies a relation between only two industries, each element of the inverse matrix reflects the interdependencies among all industries comprising the economic system. The solution to the basic static model is obtained by deriving the inverse matrix and applying it to the vector of final deliveries. Thus if one element of final deliveries changes, say final demand for cars is cut in half, the model can compute the implications for not only the output of cars but also for the outputs of all other industries, namely steel to make the cars, coal to produce the steel, energy to extract the coal, and so on. This ability to capture indirect effects is one reason for the model's popularity and for the fact that it is incorporated in virtually all empirical economic models that distinguish a sectoral level of detail, including in particular the more elaborate input-output models that will be described below.

For empirical applications, the input-output database is as important as the mathematical model. Most analyses start from tables in money units and devote a great deal of effort to making sure that the total money value of each row is equal to that of the corresponding column before deriving a coefficient matrix. By contrast, Leontief stressed the technological interpretation of each column of coefficients and urged collaboration of economists with engineers and other technological experts to project, column by column, coefficient matrices representing hypothetical changes in technologies in different industries based on information in physical units, such as that developed for the use phase of life-cycle engineering studies. The episodic collaboration of input-output economists with engineers is of long standing, but it has surged and matured dramatically in recent years. This collaboration is now being fostered by several interdisciplinary professional societies, as concerns about the environment have deepened and industrial ecologists and other environmental scientists and engineers have sought to evaluate not only the direct effects but the full, economy-wide impact of alternative technologies governing the use of energy and materials. Input-output models manipulate data in both physical and money units. They capture the direct as well as the indirect environmental impacts of alternative products or processes on the basis of the explicit representation of physical stocks and flows of energy and materials, measured in physical units, through an entire economic system. This is the reason for their value for industrial ecologists.

1.3. Physical Model and Price Model

The equation $(I - A)x = y$, with all variables measured in money units, is used in the vast majority of empirical input-output studies. However, industrial ecologists have incorporated into the input-output database data in physical units representing material flows, inputs and outputs of manufacturing processes, and product life cycles. To fully exploit the power of the model coupled with such a database for evaluating environmental impacts, it is vital to sharpen a few distinctions that often remain unclear. Goods and services were distinguished from resource inputs in Equations (1) and (2), respectively. In this section, quantities of goods and services are distinguished from their unit prices, and prices of goods and services from prices of resources and other factors of production. This framework separately tracks all of these quantities and accommodates resource commodities that are priced, either in market prices or through legislation such as carbon fees, as well as inputs or outputs that are not priced, such as fresh water or the discharge of pollutants, as seen below.

Assume in Equation (1) that each industry's output is measured in a unit appropriate for that sector, such as steel and plastics in tons, electricity in kWh, and computers and automobiles in numbers of standard units (e.g., number of computers of average capability). Some service sector output may be measured in a physical unit, such as the number of visits to a doctor's office; but other sectors may be measured in the money value of output, say dollars' worth of business services. A mixed-unit flow table accommodates variables measured in different units and can be constructed with no conceptual difficulty. In the coefficient matrix A derived from such a flow table, the ij^{th} element is equal to the ij^{th} element of the flow table divided by the jth row total. (Note that it makes no sense to calculate the jth column total in a mixed-unit table.) A mixed-unit A matrix may instead be constructed directly as columns of coefficients.

Equations (1) and (2) comprise an abbreviated form of the basic input-output model. The simplest form of the full model involves two additional equations (where $(.)'$ indicates transposition):

$p'(I - A) = v' = \pi'F$, or

$$p' = v'(I - A)^{-1} = \pi'F(I - A)^{-1} \qquad (3)$$

$$p'y = v'x = \pi'Fx \qquad (4)$$

In Equations (3) and (4), p is the vector of unit prices and v is value-added, that is, the total money value of factor inputs per unit of output. If inputs of individual factors are measured in physical units in the corresponding rows of the F matrix, such as number of workers or hectares of land per unit of sectoral output, and the unit factor prices are specified by the vector π', then value added, v', can be substituted by $\pi'F$ in Equations (3) and (4).

Equation (3) is the basic static input-output price model, and the components of the vector of unit prices are price per ton of plastic, price per computer, etc. Equation (3) shows the unit price of a good as the sum of the amounts paid out to each one of the factors of production. For a sector whose output is measured in dollars in Equation (1), for example financial services, the corresponding unit price is simply 1.0. With this equation one can compute the impact on prices of changes in technical coefficients (A) or in the quantity or price of factors (F or π'), or value-added (v'), per unit of output. Finally, Equation (4), called the income equation, is derived from Equations (1) and (3): this identity (the GDP identity) assures that the value of final deliveries is equal to total value-added (the value of all factor inputs), not only in the base-year situation for which the data have been compiled but also under scenarios where values of parameters and exogenous variables are changed.

It generally escapes notice that Equations (1) and (2) have the attributes of a quantity model when, as is most frequently the case, the outputs of all sectors and even the quantities of factor inputs are measured in money units. One component of the output vector, for example, would be the value of the output of plastic or steel, each figure being the implicit product of a quantity and a unit price, but with inadequate information to distinguish the quantity from the price. Under these circumstances, there is no perceived benefit from a separate price model: all elements of the price vector in Equation (3) would be 1.0, and the price model is therefore deemed to be trivial. This is an incorrect assessment, however, since the price model can be used to calculate changes in prices associated with changes in A or v' (e.g., improved efficiency of energy use might result in an 8% reduction in the price of a dollar's worth of business services from 1.0 to 0.92). When the variables of the quantity model are measured in physical units, the calculated prices are in money values per physical unit.

1.4. Properties of Nonnegative Matrices

To realize the full potential of input-output analysis, insight into the properties of the model given by equations (1) to (4) is needed, in particular into their economic interpretation. The properties of the input coefficients in the A matrix play a central role. Consider the case where A has been derived from a flow table by dividing the elements in each column by the appropriate total output. Clearly, if all data in the square part of the table are nonnegative, the result is a square, nonnegative matrix where all column sums are smaller than unity. But suppose A was obtained from a compilation of engineering data in physical units. Which properties does that matrix need to possess in order to satisfy the requirements of an input-output coefficient matrix? What is needed is a general theory for the equation system (1) to (4). The theory of nonnegative matrices provides the needed foundation.

Much attention has been given to conditions that guarantee that the multiplier matrix (or Leontief inverse) is strictly positive, i.e., that each element is positive. Such conditions make sense because basic economic logic would require that an increase $\Delta y > 0$ in final demand in Equation (1) should result in an increase $\Delta x > 0$ in total output. If the matrix (I

– A)$^{-1}$ were not strictly positive, this logic could be violated. We can also formulate this result differently, i.e., as an answer to the question whether Equation (1) always has a solution x > 0 for y > 0. In fact, the study of Equation (1) has led to a number of equivalent statements about A, of which:

1. $(I - A)^{-1} > 0$.

2. $(I - A)^{-1} = I + A + A^2 + A^3 + \ldots$. (That is, the series ΣA^k is convergent.)

3. All successive principal minors of $(I - A)^{-1}$ are positive.

4. There exists a choice of units such that all row sums or all column sums of A are smaller than unity.

5. A has a dominant eigenvalue λ where $0 < \lambda < 1$.

A related result is:

6. The dominant eigenvalue λ of A gets larger if one element of A is increased, and λ gets smaller if one element of A is decreased.

Statement 2 is important for distinguishing the industries contributing output in different phases of production. It says that output x = y + Ay +A(Ay) + … . So the quantity y must be produced, plus Ay which is the vector of input to produce y, etc. Statement 3 is the well-known Hawkins-Simon condition, which assures that each subsystem is productive; that is, each subgroup of industries i, j, k,… requires less input from the economic system than it produces in terms of outputs. According to statement 4, the Brauer-Solow condition, value-added in each sector is positive in coefficient matrices derived from input-output tables in (nominal) money values. That is, units for measuring physical output are such that each one costs one monetary unit (thus, if the output unit is dollars, the unit price is 1.0 by definition). Assuming that the matrix describes a viable economy, this property assures that if output is measured in any chosen physical units, there exists a set of prices such that each industry has a positive value-added (i.e., revenue left to pay for factor inputs).

The dominant eigenvalue λ is a measure of the size of the intermediate outputs produced in the economy in relation to total production. That is, λ indicates the net surplus of an economy in the sense that the larger λ (within the bounds described by statement 5), the smaller the net output. The surplus so defined can be consumed, invested for growth, devoted to environmental protection, etc. Statement 6 is useful for interpreting the role of technological change. For example, a technological innovation that reduces the need for certain intermediate inputs results in a lower dominant eigenvalue for the new coefficient matrix, leaving more surplus. Innovations that are not cost-reducing, on the other hand, will result in a larger λ. An example might be more secure disposal of hazardous wastes; see also Section 3.1. Input-output analysis can effectively identify those industries where increased technological efficiency would have a significant economy-wide impact. Thus,

λ is a kind of efficiency indicator in that of two matrices describing two different economies, the one with a larger dominant eigenvalue represents the economy that is less efficient economically although it may have other desirable features. Eigenvalues also play an important role in dynamic models, where they have an interpretation in terms of rates of growth or contraction and profit rates (see Section 3.2).

If the economy does not produce a surplus (i.e., $y = 0$ in Equation (1)), we are dealing with a closed model of the following form,

$$x = M x \qquad\qquad (5).$$

In this special case, M has a dominant eigenvalue equal to unity, and total output x is the Perron-Frobenius eigenvector of M. Solving the model for x thus means solving for this eigenvector. The solution provides only the production proportions; the scale has to be determined in other ways, such as external knowledge about the size of certain elements of x.

1.5 The Choice of Technology

Above we interpreted the ith column of the input coefficient matrices A and F as representing the technology to make good i and claimed that the coefficients represented an average technology. This interpretation allows for the existence of differences in technology among establishments in the same industry. The use of an average avoids the complication of having to distinguish products and technologies where the distinction does not add much useful information for the purposes of the analysis. In terms of the model, it means that a one-to-one relation is established between the typical commodity and an average technology for producing it.

Suppose a new technology becomes available to produce the ith good. If the industry adopted that technology, the ith columns of A and F would change accordingly. If two technologies are available for producing the ith good, the model can determine which technology is the lower-cost choice in terms of the *overall* use of factors. Equation (3) shows that the cost of factor use is equal to $v'x$ or $\pi'Fx$. If the new technology is cheaper in terms of overall factor inputs, it is more efficient than the old one and in principle will be installed. Because each price is the sum of the costs of the primary resources used directly and indirectly in its production, introducing the new technology will assure minimal cost.

The above considerations enable us to formalize the economically most efficient choice of technology as a minimization problem. Several theorems were formulated and proved in the 1950s that identify for each product the choice among several alternative technologies that minimizes the use of priced resources. It can be shown that for a particular final demand, y, there is a unique, cost-minimizing set of technologies, provided that the possibility of factor constraints is ignored. (See Section 2 for a

discussion of the more general case when one or more factors become scarce.) The Non-Substitution Theorem is the name given to the claim that this particular set of technologies will remain unchanged for any changes in final demand, providing one theoretical rationale for the fixed coefficients assumption. Once the choice of technologies has been fixed, prices are determined by Equation (3) and thus are also independent of final demand. We also see from Equation (3) that if one sector experiences an increase in efficiency (for example because of reduced coefficients in its column of A or F), all prices will fall (because all elements of the multiplier matrix will be smaller, or at least no larger, than before). In this way the model can be used to represent the economic implications of the innovation process: if a new technology becomes available, it can be represented as an additional technological option. It can be shown that, with a different objective function, say one that minimizes carbon emissions rather than costs, a different choice of technologies could be selected (see Section 3.5).

1.6. Joint Production, or Multi-Product Industries

In 1968, the United Nations issued guidelines for a new overall framework which forms the basis for today's System of National Accounts. These guidelines still assume tables in money values, and we retain that simplifying assumption for purposes of the discussion in this section. It is worth pointing out, however, that input-output tables in physical units have begun to be compiled on an experimental basis by some statistical offices, and the features described below can be readily generalized to non-monetary choices of units.

United Nations guidelines describe the compilation of input-output data in two separate tables, the make table and the use table. In the discussion until now we have assumed a square industry-by-industry input-output table, where each industry is associated with a single product. But, in fact, a business establishment may produce multiple products. Today's make-use system explicitly takes this complexity into account by making a distinction between commodities and industries, where the number of commodities and industries need not be the same. In this way, multiple outputs of any given industry can be easily accommodated in the make matrix.

Several stages, each requiring expert knowledge, can be distinguished in building the make and use tables. Data on individual firms, investments, wages, consumer behavior, imports and exports, etc., are compiled, and the corresponding totals in the make and use tables are reconciled to yield the same row or column totals. This is a time-consuming process because many sources of error, such as incorrect data, errors of classification, or cases that are ambiguous to classify must be resolved. A separate challenge is selecting appropriate classifications and definitions for commodities and industries.

While the make-use system has greater information content, its use adds another step for the analyst: constructing a square input-output table from the two rectangular ones. Methods differ in the assumptions regarding the technologies used to produce a given commodity in different industries. Well-known are the so-called commodity technology

and industry technology assumptions. The former assumes a single technology for producing a given product regardless of the sector that produces it, while the latter preserves institutional distinctions and assumes a common technology for all the products produced in a given sector. Each method has its proponents and may be better suited for particular empirical cases; both are internally consistent logically and consistent with the System of National Accounts

1.7. Approaches to Analysis

There are three dominant approaches to analysis using a basic static input-output model: multiplier analysis, decomposition analysis, and scenario analysis. The first two start from an input-output flow table for a past period of time and are used to reveal underlying relationships among sectors. Scenarios, by contrast, are generally used to evaluate the consequences of alternative assumptions, in particular about the future.

1.7.1. Multipliers

Multiplier analysis is widely used to analyze the impacts of shifts in final demand on total output or total factor use. If (in Equation (1)) final demand y changes to $y + \Delta y$, where the individual elements of Δy can be positive, negative or zero, we have:

$$(x + \Delta x) = A(x + \Delta x) + (y + \Delta y),$$

which is interpreted as the summation of $x = Ax + y$ (which reflects the initial situation) and $\Delta x = A\Delta x + \Delta y$. Solving for Δx yields

$$\Delta x = (I - A)^{-1}\Delta y.$$

We point out here the economic interpretation of the results of this mathematical manipulation. For example, $\Delta y > 0$ implies $\Delta x > 0$, because the matrix $(I - A)^{-1} > 0$ (from Section 1.4). The economy's interdependencies thus are reflected in the fact that if final demand for only one commodity changes, all sectors will experience a change in output. For example, if $\Delta y_i > 0$, while final demand for all other commodities remains unchanged, we have $\Delta x = a^{\cdot i} \Delta y_i$ where $a^{\cdot i}$ represents the ith column of the multiplier matrix. That is, all industries will have to increase their production, and corresponding increases in factor use can be obtained straightforwardly.

The concept of key sectors plays an important role in applied research. Key sectors can be defined in several ways. Some well-known methods focus on the relative size of elements in the multiplier matrix. If, e.g., element a^{ij} of this inverse is relatively large, this means that a decrease in final demand for commodity j will have a significant downward pressure on the output of industry i, and, consequently, on employment or carbon emissions in that industry. The foundation of much of economic stimulation policy is based on this type of observation. The effect can be particularly powerful in instances

where economic interconnections are strong, such as in certain geographic regions or among tightly integrated sectors.

1.7.2. Decomposition Analysis

Structural decomposition techniques are used to explain the change in one particular variable in terms of changes in other variables. For example, we may wish to explain output growth between time periods t_1 and t_2 by decomposing it into quantity changes and price changes, each of which in turn may be broken down further into sectoral detail. A variety of decomposition methods have been developed to explain a change in sectoral output by distinguishing the effects of changes in technical coefficients (between A_1 and F_1 at t_1 and A_2 and F_2 at t_2) from those of changes in final demand (y_1 and y_2): private consumption, government expenditures, exports and imports, and capital investments. Technological change, in turn, may be a response to shifts in output composition or level, or in its distribution over the final demand categories (as reflected, for example, in changes in international competitive position). Applications usually are carried out using input-output tables in money units measured in constant prices, in which case changes in output reflect changes in quantities. Such techniques have also been used to decompose multipliers into distinct components and, by extension, the decomposition approach has been applied to analyzing the impacts of energy conservation policies and for gaining insight into the impact of privatization efforts.

For an example, take the price equation, Equation (3), and let us say that we are interested in explaining price changes in terms of changes in the payments to factors per unit of output. Simplifying the notation by dropping time subscripts, we have at time t_1:

$$p' = v'(I - A)^{-1}$$

Taking the differences from the point of view of time t_2, we obtain:

$$\Delta p' = v'\Delta(I - A)^{-1} + (\Delta v')(I - A)^{-1} + (\Delta v')\Delta(I - A)^{-1}$$

(If the changes in the variable values approach zero, the standard results from the product rule of calculus are obtained.) The price changes are expressed as the change in each variable multiplied by the base level value (i.e., the value at time t_1) of the other variable, plus an interaction term in the changes in both variables. The interaction term requires a subjective decision about how to allocate it among its component terms.

1.7.3. Scenario Analysis

As we have seen, an A matrix that satisfies the Hawkins-Simon condition (Section 1.4) can be used to derive essential statistics, namely multipliers and eigenvalues, that describe the structure of the economy in question in ways that cannot be observed directly. It can also be used to decompose a change in one variable into changes in other,

explanatory variables. A different analysis, however, is to assess the impacts on output, factor use, pollutant generation, and prices of hypothetical scenarios describing changes that could take place in that economy in the future. A scenario can be very simple, for example specifying a 20% reduction in energy inputs to all sectors due to efficiency improvements. More elaborate scenarios are more concrete and detailed; for example, one could specify alternative technologies for the generation of electric power. Such scenarios are often based on case studies, such as engineering process analyses that yield several columns of coefficients in A and F corresponding to alternative technologies. Innumerable studies of both types have been carried out, the former mainly by economists and the latter mainly by industrial ecologists.

In the earliest scenario studies, voluminous results were reported, in the form of many pages of tables and figures. Today the questions are often more focused and the reported results more concise. The definition of a scenario identifies the variables and parameters about which assumptions are being made as well as the endogenous solution variables that will be used to evaluate the scenario. One structured approach to scenario analysis is to pose a question in the form of a hypothesis and then design one or more scenarios to test it.

2. Beyond the Basic Static Model

Like every model, the basic static input-output model has a number of limitations. Sometimes these can be overcome while remaining within the basic static framework and continuing to rely on a matrix inversion to capture interdependencies. That is the case for the analysis of environmental problems that will be described in Section 3.1. Other extensions require changes to the model structure, as exemplified by the subsequent examples. These extensions involve other solution concepts than the inverse of a single matrix.

2.1. Making Exogenous Variables Endogenous

Whether the basic static model is used to evaluate scenarios or to derive multipliers or eigenvalues from the inverse matrix, its crucial limitation is that many key variables need to be specified exogenously: all categories of final demand in the physical model and all categories of factor costs in the price model (i.e., the righthand sides of Equations (1) and (3)). For past years, the values of all exogenous variables are known, so this is not a problem for computing eigenvalues, multipliers, or decompositions. However, they obviously are not known for the future. Furthermore, while changes in technological assumptions influence the prices of goods (i.e., changes in A or F lead to changes in p' by Equation (3)), there is no endogenous feedback from changes in prices to changes in physical quantities (i.e., changes in π' or p' do not affect the choice of technologies that are represented in A and F or the value of x) or from changes in demand to prices (i.e., y to p'). These feedbacks are crucially important in real world situations, excluded by the Non-Substitution Theorem, where limited factor availability is common. The following sections describe ways in which the conceptual reach of the basic model has been

extended to overcome some of these limitations while remaining an input-output model. Often the input-output inverse (essentially the monetized form of Equation (1)) is incorporated into models that are based on assumptions about economic interdependency that are different from those underlying an input-output model. This is the case for computable general equilibrium (CGE) models, which rely extensively on a variety of elasticity parameters that are not described in this paper.

Scenario analysis with the basic static input-output model requires the analyst to make many exogenous assumptions. The collaboration with engineers and other technological specialists provides a firm basis for some of these assumptions, namely those about changes in technologies as reflected in the A and F matrices. From an input-output point of view, use of expert knowledge in the form of exogenous assumptions about technological options is superior to relying on formal methods to represent behaviors that make technological change endogenous.

However, the input-output economist must make assumptions not only about technological alternatives but also about changes in the levels of various categories of final demand and in factor prices, and these assumptions must be consistent both with each other and with the technological projections. Extended input-output models make some of these relationships endogenous by introducing equations to increase the closure of the model, one or a few variables at a time. (The model would be completely closed if there were no exogenous variables at all. This situation can be interpreted as the absence of economic surplus, see Section 1.4, or the case where all uses for the surplus have been made endogenous.) New variables and parameters are introduced, broadening the scope for hypotheses and scenarios, but also increasing data requirements and making more challenging the economic interpretation of a wider set of numerical results.

Multipliers and eigenvalues can be defined only in the context of a linear input-output model, that is, in terms of a square matrix transforming a vector of final demand into a vector of outputs (Equation (1)) and transforming the vector of factor costs into the vector of prices (Equation (3)). The linearity of the model means that a given percentage change in the level of final demand results in the same percentage change in the level of output (and analogously for factor costs and prices of goods). Analysis based on multipliers and eigenvalues is appealing because – unlike scenario analysis -- it specifies explicit variables (namely, the multipliers and the eigenvalues) to be quantified as the outcome of an analysis, and the formulas for their computation are clearly specified. By contrast, the choice of variables that constitute the results of a scenario analysis is completely open-ended. For these reasons, the most common approach to extending the basic model is to make exogenous variables endogenous by adding for each a row and a column thereby increasing the dimension of the A matrix. Then the problem remains linear and the algebra of the basic static model can still be applied, but to a coefficient matrix of larger dimension. The classic example of this approach is the analysis of the so-called Social Accounting Matrix, a matrix that in particular associates the row of labor coefficients in money values with the column of outlays of this income for consumption goods and services. It is said to be closed for households, since it makes household incomes and outlays endogenous. The Social Accounting Matrix is treated like an input-

output coefficient matrix, A, in that its inverse is computed to arrive at a solution. However, closure for investment (in the form of a dynamic model) and trade (a world model) move beyond the familiar linear model and are solved without recourse to a Leontief inverse matrix.

2.2. Simplifications and Transparency

Important extensions of the basic static input-output model include a model closed for households, a dynamic model closed for investment, a model to analyze pollution generation and the impact of abatement technologies, and a multi-regional model of the world economy that is closed for trade. It is natural to ask why these extensions are described one at a time rather than cumulatively. Why not study abatement prospects with a dynamic world model closed for households? The answer is that each such combination of closures requires more economic theory-building than does the collection of individual closures. For example, a dynamic model with closure for households requires the endogenous representation of household purchases and use of capital goods; a dynamic world model requires the endogenous determination of stocks and flows of not only domestic but also foreign direct investment. Solutions to these combined challenges have not been attempted within an input-output framework and remain a more or less distant goal.

A popular approach to modeling is to introduce the many kinds of real-world complexities, such as tariffs in a model of international trade, which can obviously improve the fit between calculated quantities and observed ones. The models described here are expressed in terms of fundamental concepts only in order to retain their transparency, that is, the demonstrably clear logical links between assumptions and conclusions. Complexities (such as exogenous tariffs) can readily be incorporated in empirical applications.

A model is best described in terms of the simplest economic logic that can do the job. Thus, it may be convenient to use the monetized concept of a value-added vector, v, rather than the richer representation of $F'\pi$ in cases where the analysis is not focused on physical quantities or unit prices of specific factor inputs. Likewise in the case of the dynamic input-output model, attention will focus on investment to expand production capacity: other phenomena, such as the logic behind investment for replacement of capacity already in place, will for now be ignored.

2.3. What is Distinctive about an Input-Output Model?

The model extensions described below are designated as input-output models, while the integration of input-output matrices into, say, a computable general equilibrium model does not make the latter into an input-output model. What then distinguishes an input-output model from other models making use of input-output matrices?

The distinctive feature of an input-output model extended by the closure for particular variables is that the closure is achieved at the economy-wide level and the sectoral level simultaneously rather than consecutively, if at all. Other models solve first for an economy-wide solution and then use an input-output matrix to disaggregate macro-level results to the sectoral level. However, the sectoral distribution affects the aggregate value because of both direct and indirect effects. To have a chance of achieving consistency between macro and sectoral levels, the solution process would need to be iterated, and there are no *a priori* grounds for assuming convergence in any particular model.

The conceptual consistency assured by input-output models becomes clearer if specific instances are examined. For two cases – the closure for households and the dynamic model -- the input-output closure is briefly described below, showing how aggregate levels and sectoral implications are simultaneously determined. The dynamic model is described in more detail in Section 3.

2.3.1. Closure for Households

Assume for simplicity that labor inputs are measured in money value. Then the level of labor income is endogenous in the basic static input-output model in that, when the level of consumption increases or decreases, or its composition changes, earnings of labor in each sector change accordingly. However, since the level of consumption is exogenous, when labor requirements increase or decrease due to changes in coefficients or in levels of final demand, outlays for consumption do not automatically change to reflect the changes in purchasing power. The level of consumption changes only if it is adjusted exogenously as part of a scenario.

In the case of the typical closure for households, the labor row and consumption column are made interdependent so that the amount of income governs consumption outlays, and the amount of consumption is a major determinant of the demand for labor and associated income. This closure assures consistency among labor requirements, labor income, output, and the level and composition of consumption not only for the economy as a whole but also on a sectoral basis.

2.3.2. Closure for Investment

When the level of investment increases or decreases or its composition changes exogenously in the basic static model, output changes accordingly and requirements for capital (and other factors) also change (Equation (2)). However, there is no assurance that the exogenous the level of investment is consistent with the endogenous changes in outputs, nor do calculated prices change to reflect the changed costs of capital investment (Equation (3)).

A dynamic input-output model achieves closure by assuring that each sector purchases in a given time period enough capital goods to assure adequate capacity for its own

anticipated future production and produces enough capital goods to satisfy the current investment orders (for future additions to their capacity) of other sectors. In other words, inter-temporal consistency is assured for both quantities and prices at a sectoral level. It also assures that prices cover a return on the capital stock adequate to cover the investment in additional capital goods. A more detailed discussion, that also indicates the particular challenges faced by dynamic models, is provided in Section 3.2.

3. Major Model Extensions

In this section, we describe three major extensions of the basic input-output model. These extensions address the generation of pollution and policies to combat it, capital investments in dynamic models, and multi-regional models, including an input-output model of the world economy.

3.1. Environmental challenges

Input-output models can represent pollution and ways of reducing pollution and assigning responsibility for it. Pollution is defined here as the production of undesired and noxious products simultaneously with the production of economically valued ones. An example is the emission of carbon oxides into the atmosphere as a by-product of electricity generation. Input-output analysis establishes a direct relation between either the output level, or the amount of fuel combusted, in the emitting sector and the quantity of carbon emissions generated. For each sector pollutant-specific emission coefficients represent the quantity of pollutant produced per unit of output (or of fuel input) of the polluting industries. For example, let c be the vector of carbon coefficients (this could include zeroes for sectors that emit only insignificant amounts of carbon). If x is the vector of total output, the product $C = c´x$ represents the total volume of carbon.

Emission coefficients for specific sectors and technologies have been estimated for many pollutants such as nitrogen, sulfur, phosphorus, the heavy metals, and their compounds. These coefficients may be assembled into one matrix of emission coefficients, a row for each pollutant. Writing A_{11} for the standard coefficient matrix (A in Equation (1)) and A_{21} for the emission coefficient matrix:

$$(I - A_{11}) x_1 = y_1$$

and

$$A_{21}x_1 = y_2$$

where x_1 and y_1 replace x and y in Equation (1), and y_2 stands for the total amount of pollution entering the environment. Notice that the emission coefficient matrix, A_{21}, pre-multiplies the output vector, just like the matrix of factor inputs, F, in Equation 2.

Let us assume that environmental policy requires limiting emissions through commodity-specific abatement technologies, such as equipment for carbon capture. These measures can be incorporated into the technology for the polluting industry, or they may be treated as separate industries. In the latter case, let A_{12} represent the input requirements of the abatement activities per unit of pollution eliminated, x_2 the quantities of pollutants eliminated by the abatement activities, and A_{22} the matrix of emissions still remaining after treatment by the abatement activities. The model now reads as follows:

$$(I - A_{11})\, x_1 - A_{12}x_2 =\ y_1 \qquad (6)$$

and

$$A_{21}x_1 - (I - A_{22})\, x_2 = y_2 \qquad (7)$$

where y_2 measures the amount of pollution still reaching the environment. Given exogenous values for y_1 and y_2, the model determines x_1 and x_2 using the multiplier matrix of this extended system. The corresponding prices are calculated via the dual to the quantity system.

The above system is a direct extension of the basic static input-output model. However, it imposes additional non-negativity and other constraints. (For example, the quantity of tolerated pollutants can not be larger than the quantity produced.) If the same pollutant can be generated by several different industries, one can introduce a distribution mechanism that allocates specific emission targets to each polluting industry.

A scenario could require that all industries lower their emissions by the same percentage, although it may be desired to place tighter constraints on certain industries than others since those with a strong market position can pass on any additional costs to their customers through increases in prices. The impacts of different arrangements can be explored in this framework. Also, the effects of a market for emissions can be analyzed, solving for changes in quantities of emissions, outputs and resulting prices for given final demand.

Sometimes a situation exists in which cleaner technologies are available but more costly than alternatives. In this case incentives like targeted taxes and subsidies may be required to make their adoption cost-effective. The model described above can be used to calculate the cost and price implications under alternative cost-distribution schemes. The impacts of technological improvements similarly can be investigated using the model: if they result in lower pollution, abatement costs also will be lower. The obtained figures can be compared with the development and introduction costs of the newer but cleaner technologies.

Societies ultimately need to decide between placing responsibility for reducing pollution with consumers or producers. Since household consumption is considered the ultimate goal of economic production, it is logical for environmental policies to include incentives for changing consumption behavior, for example via a system of consumer taxes and

subsidies. Such policies introduce issues of 'fair' sharing of burdens between rich and poor segments of society, and between rich and poor societies, not to mention the challenges of political acceptability for governing bodies.   These distributions are readily represented at the level of detail of an input-output analysis.

Pollution and pollution reduction can be represented in dynamic input-output models and in multi-regional models including models of the world economy in similar ways. Increasingly, input-output models of the economy are linked with models of parts of the ecological system in a single integrated framework for studying a particular set of problems. Examples are studies of the linked economic and environmental attributes of fisheries and forests, or the integration of an input-output model with a physical model of land use or hydrological flows. These studies attempt a partial closure of the combined model for specific material subsystems, such as the recycling of materials or the economic use by one industry of another industry's waste products. This process is called "closing the loops" in industrial ecology.   While economic theory provides extensive guidance about the role of factors in production and trade, the relationships between factor inputs and waste outputs are not yet fully theorized. This closing of the loops is an important step in that direction.

3.2. Dynamic Input-Output Models

Investment in capital goods, in particular those embodying new technologies, is the principal mechanism for achieving changes in the structure of an economy including the adoption of less-polluting techniques. Representing investment requires the introduction of an explicit time dimension into the economic model. A sector's investment decisions at one period of time depend on expectations of future demand for its product. One sector's current investment calls for a variety of goods and services to be produced over some interval of years by other sectors but will add to the former's production capacity only with a lag of the entire time interval. A dynamic input-output model must track these intertemporal and intersectoral relationships and assure their consistency. To be useful for evaluating scenarios about sustainable development, the model must neither assume growth nor preclude it, both for individual sectors and for the economy as a whole. That is, sectors must be able to contract, to maintain their capacity, or to expand.

The first formulation of the dynamic input-output model took the following closed form (that is, with no exogenous final demand), written here in terms of the matrix equation for a single time period.

$$x_t = A\,x_t + B\,(x_{t+1} - x_t). \qquad\qquad (8)$$

or, solving for x_{t+1}:

$$x_{t+1} = [I + B^{-1}\,(I - A)]\,x_t \qquad\qquad (9)$$

Output in period in period x_{t+1} thus is obtained by a complete reinvestment of x_t, the previous period's output.

While a more elaborated dynamic model is described below, results obtained with Equation (8) provide a useful benchmark for comparison by revealing the maximal possible rate of proportional growth, or the minimal rate of contraction, available to that economy. This rate of growth (or contraction) can be calculated from the A and B matrices.

Balanced growth means that all sectors grow at the same rate, or

$$x_{t+1} = (1 + \gamma)x_t$$

where $\gamma > 1$ in the case of growth. Substitution in the closed form results in:

$$x_t = Ax_t + B[(1 + \gamma)x_t - x_t]$$

$$= Ax_t + \gamma Bx_t$$

or, rewriting,

$$(1/\gamma)x_t = [(I - A)^{-1}B]x_t \qquad (10)$$

Denoting as μ the dominant eigenvalue of matrix $(I - A)^{-1}B$, we thus have $1/\gamma = \mu$, or $\gamma = 1/\mu$.

The accompanying price model can most easily be understood from an individual sector's perspective. If b_j is the vector of capital goods needed to produce one additional unit of good j, the value of these capital goods is $p'_t b_j$, where p'_t is the vector of prices at the beginning of period t. If the sector produces one unit of j, and if bills are paid at the end of the period, i.e., in prices of period t+1, the sector's intertemporal price relations are described as:

$$p'_{jt+1} - p'_{t+1}a_j + p'_{t+1}b_j = (1 + r) \, p'_t b_j \qquad (11)$$

where a_j stands for the jth column of the A matrix; the righthand side can be called the opportunity cost of depositing the money value of the stock at a bank at the interest rate r. For the economy as a whole, we then have:

$$p'_{t+1}[B + I - A] = (1+r) \, p'_t B$$

or

$$p'_{t+1} = (1+r) \, p'_t[I + (I - A)B^{-1}]^{-1}. \qquad (12)$$

We should note that equation (8) is an extension of Equation (3): if prices are p′ for all t, we find p′ = p′A + rp′B, where the term rp′B has replaced the value-added term. Note that we have abstracted from maintenance and replacement of capital as well as from all non-capital factors of production.

The shortcomings of this dynamic model for empirical analysis include the extent and nature of its instability. It turns out, by the Dual Stability Theorem, that the output and price models diverge, meaning that if one is stable, the other is unstable. This occurs because the matrices $[I + B^{-1}(I - A)]$ and $[I + (I - A)B^{-1}]^{-1}$ have different properties: the eigenvalues of each matrix are reciprocals of the eigenvalues of the other. In economic terms the instability is explained by the fact that the outputs and prices of one period are the inputs for the next one.

One approach to making the dynamic model better equipped for empirical analysis was to turn it into the form of the basic static model and compute what was, for obvious reasons, called the dynamic inverse. The system-wide coefficient matrix consists of blocks of matrices where $R = (I - A + B)^{-1}B$ determines its properties. The entries correspond to different time periods in the time horizon of interest, and the entire matrix pre-multiplies the vector x of sectoral outputs in each time period. This model does not suffer from instability problems like the other variants of the dynamic model and attracted substantial interest in the 1960s and 1970s, when its mathematical properties were widely explored. However, it failed to resolve the major limitations of the model for empirical analysis.

To overcome these limitations, another dynamic model was introduced. The new model requires an additional variable: each sector's production capacity, or the maximum output it can economically produce with a given capital stock. The vector of base-year rates of capacity utilization needs to be exogenously specified as part of the initial conditions. In this model a sector invests in expanding its capacity if and only if two conditions are met: its output is growing, and its capacity is fully utilized. If either condition is not met, no expansion investment takes place. Closure for investment and profits thus is achieved by introducing this non-linearity: as a consequence, doubling demand in a given time period will in general not result in doubling output because production of goods for investment is dependent upon not only final demand but also the sectoral rates of capacity utilization. The model allows for changes over time in the coefficient matrices, indicated by the subscript t. The simplest version of this model, assuming only a one-year time lag for all capital goods, can be written as follows:

$x_{t} = A_t x_t + B_{t+1} O_{t+1} + y_t$ or

$$(I - A_t)x_t - B_{t+1} O_{t+1} = y_t \qquad (13)$$

with

$$O_{t+1} = \max [0, c^*_{t+1} - c_t] \qquad (14)$$

$$c_{t+1} = c_t + O_{t+1} \qquad (15)$$

where (dropping the time subscript, t, where no confusion results): A, B, x, and y are defined as before, c stands for capacity, and O for desired addition to capacity. The vector c*, the desired capacity for time t+1, is projected in a side calculation as a moving average of recent past rates of growth of output. The initial conditions include values for c_{t0}, c^*_{t0+1}, and O_{t0+1}. The full output model includes additional equations representing factor inputs and a price equation like Equation (12) but with time-specific matrices of parameters.

Equation (13) is similar to Equation (8), except that now the so-called accelerator term in B_{t+1} may be zero and, even when it is non-zero, it is not expressed in terms of output, x. This fact frees the solution from the knife's-edge property of Equation (8), where in each period output had to be exactly that quantity that assured the full utilization of the capital stock.

The dynamic model described by Equations (13) – (15) and the corresponding price model can be used to analyze scenarios based on alternative assumptions about the exogenous variables (y) and technical coefficients, describing not only the current-account inputs (A and F) but also initial rates of capacity utilization, capital coefficients (B), and the interest rate. The model is used to calculate outputs and prices year by year over a particular stretch of time, with the patterns of changes in output defining investment cycles at the sectoral level and economy-wide.

For alternative scenarios, A, F, and B need to be projected simultaneously. The collaboration of economists and engineers is crucial to be sure that the scenario is based on technological information and also treats production capacity, the expansion of the capital stock, and the earnings of capital consistently from an economic point of view.

The dynamic model makes investment and profits endogenous in ways that can not be accommodated satisfactorily by the basic static input-output model but require the introduction of additional variables and parameters and a non-linearity in the relationship between demand and output. Now, output is dependent not only on technology and final demand but also on the rate of utilization of the capital stock. For this reason, the solution of this system cannot be deduced from a single system-wide matrix inverse and its multipliers or eigenvalues.

3.3. Regional and Multi-Regional Input-Output Models

The models discussed above treat a national economy as spatially homogeneous. That is, they abstract from differences with a geographical dimension. The usefulness of such geographically averaged attributes may sometimes be limited when there are significant differences between parts of the region under consideration, say a country. These may be climatic differences or differences in population attributes, natural resource endowments, level of technological development, and so on. A shift in national final demand therefore may have very different impacts in the various regions, and a change in one region may

have important impacts in other regions. In many cases, 'regions' are well-defined entities with distinct administrative or (sometimes former) political borders. Well-known examples are Northern and Southern Italy, the Western and Eastern states of what is now the Federal Republic of Germany, the 50 states of the USA, or the Western, Central, and Eastern regions of China. There is a long and well-developed tradition of using input-output models for regional economic analysis.

Typically a regional economist wishes to study not a single region but the relations among neighboring regions, in particular how events in one region impact on other regions. For this purpose each region requires a table describing 1) its within-region, or intraregional, production, and one or more tables describing2) its export of commodities to industries and final use in the other regions and 3) its imports from the other regions. While compiling reasonably accurate and up-to-date regional tables is already problematic, the greatest constraint is in the limited availability of data on interregional commodity flows. A number of methods have been devised to estimate these flows to calculate the parameters called trade coefficients. (While the terms "exports" and "imports" are usually reserved for trade that crosses international borders, we have used them to refer to all interregional flows.)

Two prominent approaches to regional analysis are interregional input-output analysis (IRIO) and multiregional input-output analysis (MRIO). A product flow in the IRIO model requires four indices to distinguish the sectors of origin and destination as well as the regions of origin and destination. In the case of two regions, these data are organized in four tables, two registering the intraregional transactions among industries within each region and the other two describing the interregional trade transactions from a sector in one region to a sector in the other. (In the case of k regions, k^2 tables are required.) The intraregional tables are square, while the tables representing the trade flows may be rectangular if the product classifications are not the same in all regions. Intraregional input coefficients are obtained in the standard way; that is, each input flow is divided by the total output of the sector using that input. The interregional trade coefficients are obtained similarly: the denominator is the total output of the using sector in the receiving region. A multiplier matrix for the entire inter-regional system is also obtained in the standard way, and the system yields a unique solution, given exogenous final demand of all regions, for each region's output. The main feature of this approach is that each region's flow table (before calculating the corresponding A matrix) is divided into several portions, one describing the output produced in the region that stays in the region and the others describing the products produced in the region that become inputs to the other regions. For a two-commodity, two-region model, we have:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where A_{11} and A_{22} are the intraregional coefficient matrices and A_{12} and A_{21} the matrices of trade coefficients; x_1 and x_2 are commodity outputs in both regions, and y_1 and y_2 are final demand. Often y_1 and y_2 are subdivided to distinguish the regional origin of imported consumption goods.

It is generally recognized that the off-diagonal A_{ij} matrices require data that are hard to come by and that in any case are not a good choice of parameters since they are considerably more volatile than the coefficients of an A matrix of technological coefficients. For these reasons the MRIO model was formulated in terms of regional technical input coefficients, A: these are the sum of the intraregional input coefficients for a region and the matrices representing the flows from other regions in the IRIO framework. Information on inter-regional flows is supplied by diagonal matrices of coefficients, C_{km} , where each diagonal element represents the amount of commodity i coming from region k to region m as a fraction of the total supply (i.e., intermediate and final) of good i in region m. (The sum of corresponding diagonal elements adds to 1.0, accounting for the full supply.) The same supply structure also is postulated for final demand. We have, for an n commodity, two-region model:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

or, in compact notation

$$x = CAx + Cy$$

That is,

$$x = (I - CA)^{-1} Cy.$$

Here x_1 and x_2 are output vectors in regions 1 and 2, and y_1 and y_2 are final demand. The model is solved like the basic static model by inverting the system matrix.

An input-output table compiled at the regional level could be more precise than a national table because regional firms are likely to use a narrower range of technologies than the 'average' technology in national tables. For example, motor vehicle factories in one region may produce only small personal cars while another region may specialize in military vehicles. Nonetheless, there are severe limits to acquiring regional input-output data because there is much less systematic compilation at this level of geography. Another problem in regional analysis is that the fixed coefficient assumption may be more problematic than at the national level; if inputs depend on only a single firm, or on a limited number of firms, they may be volatile even if technologies do not change. Another constraint is confidentiality; official data will not be released in the case of only one or a small number of firms.

Other extensions of the basic input-output model are available nowadays for multi-regional analysis, with the particular methodology customized for addressing a particular set of issues. In what follows, we will use the general term "multi-regional" for all the variants of models with exchanges among several regions.

3.4. World Input-Output Models

Economists first created mathematical models of the world economy in the 1970's in the attempt to understand the likely implications of the sharp increase in oil prices. A global framework was needed for this analysis because petroleum reserves are concentrated in certain parts of the world while used most intensively in other countries, and the magnitudes of the effects on diverse economies could be great enough to lead to substantial feedback effects throughout the global economic system. The importance of the problem called forth the resources to fund an economic research program of unprecedented ambition. The effort would require not only suitable theoretical and methodological frameworks but also a database with comparable definitions, conventions, level of detail, and information content for all parts of the world. The early world models and databases, including the first input-output model of the world economy, were funded through the United Nations.

The first world input-output model took the particularly simple form of the basic static model but with many regions. The input-output matrix in this case was a bordered, block-diagonal matrix having the A matrices for all regions as blocks down the diagonal and matrices of parameters governing exports and imports as borders in the right-hand columns and bottom rows. While the matrix included many additional equations, total world output was basically computed as a linear function of exogenous final demand, as in a multi-regional input-output model. Each region's imports and exports were made endogenous by introducing two sets of (exogenous) parameters: import-to-output ratios for imports and share-of-world-export ratios for exports for each region and sector.

The project was bold in several ways. For one thing, it marked the first time an input-output database with dozens of sectors was assembled simultaneously for each of 15 regions comprising the entire world economy. Requiring large amounts of both data storage and computational capabilities for that time, the model consisted of thousands of individual equations that calculated quantities of goods and services produced, consumed and traded in each region for the period 1970 through 2000 under alternative scenario assumptions. It also calculated land use and employment, estimated investment requirements, and calculated prices, while assuring that all these quantities were consistent.

There is a clear overlap between the multi-regional models of sub-national regions of Section 3.3 and a model of exchange among countries, or regions larger than one country, in a model of the world economy. The strategy underlying the parameterization of imports and exports is similar in both cases, but the world model has been used to analyze alternative scenarios about the future while the models of Section 3.3 are typically used to compute multipliers that quantify the impact on one economy of an exogenous change experienced in one economy on another economy.

3.5. A World Input-Output Model Based on Comparative Costs with Factor Constraints

In the early 21st century there is renewed interest, following two decades of relatively little activity, in models of the world economy. Access to petroleum remains an important concern as it was in the 1970's, but today strategies for achieving sustainable development, vague though that term may be, are also a high priority. Such strategies must take account of the changing international division of labor associated with globalization, the growing demand for resources, both apparently abundant and obviously scarce, and the environmental implications of both production and consumption. A global framework that can reach down to a sectoral level of detail is needed for this analysis. New sources of financial support are now available for both model development and data collection, and the level of modeling activity increases in parallel.

One use of such models is to evaluate the feasibility and the desirability of alternative scenarios about the future, in particular regarding the energy use and greenhouse gas emissions associated with different assumptions about resource availability, technologies, and consumption requirements. Relevant measures include both the energy actually used in each economy, including that embodied in its exports, and the energy required for its consumption, independent of where the inputs to its consumption were produced.

A model of the world economy at a moderate level of detail is intrinsically more complex, and naturally requires a more extensive database, than a one-region model. A crucial advance was to replace the trade parameters and related equations by incorporating the logic of comparative advantage into the production equations, whether of the multi-regional models described in Section 3.3 or the first world model. According to the theory of comparative advantage, a country will produce and export goods and services for which it is the relatively low-cost producer until it runs out of one or more factor inputs, at which time the next relatively lowest-cost producer will start production. The world price will be set by the highest-cost producer that actually enters into production, with extraordinary profits, called economic rents, being earned by the lower-cost producers. Differences in cost structures among potential trade partners are explained by differences in resource endowments, techniques of production, and consumption patterns, and are of course subject to influence by corporate strategies and government policies. These variables and parameters are precisely the building blocks for describing the structure of each individual economy in an input-output model.

A recent improvement on the world trade model formulates it as a linear program that optimizes an objective subject to constraints. This model has multiple practical advantages: it obviates the need to project trade parameters and vastly reduces the number of equations and the amount of computation. Most important, however, is the fact that it provides a theoretical basis, that of the theory of comparative advantage, for interpreting the numerical results corresponding to different scenarios. One version of this model is shown in Section 3.6 in Equations (16) through (21). The primal corresponds to the physical model, the dual to the price model, and the equality of the primal and dual objective functions plays the role of the income equation in the one-

region model. The variables that have not yet been introduced in the one-region models are described below.

A key feature of this world model is the role played by factor endowments, f, as constraints on the availability of resource inputs, Fx ≤ f. Not only is a region's output no longer a linear function of its own final deliveries, but world output is no longer a linear function of total final demand. This is true because the determination of relatively low-cost producers depends not only on their techniques of production but also the availability of production factors, and the amount oif output they are capable of producing is limited by factor endowments. Goods prices, likewise, are no longer a linear function of factor costs because fully-utilized factors earn, as they do in reality, scarcity rents; but no scarcity rents are earned on factors with unused capacity.

The Non-Substitution Theorem (Section 1.5) provided a theoretical rational for the fixed coefficients of the basic static input-output model, but on the clearly unrealistic assumption of unlimited availability of factors of production. The world trade model relaxes that assumption. Since international trade is a vital part of the contemporary world economy, a region's consumption is not limited by its own endowment of factors, since other regions' factors are embodied in its imports. Thus the appropriate context for introducing factor constraints is in a world model, where factor endowments are constrained at the regional level, as in Equation (18) below.

The new model introduces added realism into the calculations. Because the model is intended for analyzing scenarios about sustainable development, the primal has been implemented to minimize factor use for exogenously given final deliveries. Other objective functions could be used instead: for example, consumption could be maximized at given factor use in the spirit of a growth model. The objective function could also require the minimization of pollutants, say carbon emissions. In applications that have been carried out recently, economic and environmental objectives were specified in an alternative objective function and a trade-off curve computed to quantify the costs of progressively deeper reductions in carbon emissions.

3.6. Bilateral Trade and Geographic Interdependence

A defining characteristic of any input-output model is its ability to calculate indirect production requirements: a solution can trace an item of final demand, say automobiles delivered to final consumers, back to the engines, bodies and tires from which they are assembled and from there can quantify the steel and electronic components, as well as the coal and iron ore, not to mention the labor and capital, used at each stage of production and in total. In the case of the one-region model, these different stages represent production taking place in different sectors, such as steel production or coal mining. The world input-output model could in principle track these stages of production back to the specific economies where the production actually takes place. Thus the world model makes it possible to identify the geographic sources of inputs and reflect the input structures used in their production or extraction. The fundamental requirement for

satisfying this objective is for the model to determine bilateral trade flows. This is achieved by internalizing into the cost comparisons the transport costs associated with each region's potential trade partners, depending on the transport mode and quality of infrastructure, commodity carried, and actual physical distance between origin and destination regions. The information is contained in a transportation matrix for each pair of regions, with the cost of transporting imports charged to the receiving region. The parameters in the transportation matrix are far more reliably projected to the future than import-to-output coefficients or export shares. The input structures of the transport sectors are subject to the same kinds of technological changes as all other sectors. As for the effective distances between regions associated with a particular model of transport and the bulk of a particular commodity transported, these do change, but slowly and in well-understood ways. With this framework it is also possible to quantify the energy expended for international transport of a region's imports.

The world trade model with bilateral trade is shown in Equations (16) through (21). A, F, x, y, p, and π are region-specific with the same interpretations as in a one-region model. The new variables are e_{ij}, representing exports from region i to region j (and equal to imports to j from i), and r, representing the extra profits, or scarcity rents, earned on factor endowments that are fully utilized. The basic world model linear program needs no additional parameters, but the version below, when furnished with the matrix T_{ij} designating ton kilometers of transport from i to j distinguished by individual goods and means of transport, makes it possible to calculate bilateral trade flows.'''$\Sigma\Sigma\Sigma$

The primal problem is:

$$\min \quad z = \sum_i \partial_i' F_i x_i \tag{16}$$

$$(I - A_i)x_i - \sum_{j\neq i} e_{ij} + \sum_{j\neq i}(I - T_{ji})e_{ji} \geq y_i \quad \forall i \tag{17}$$

$$F_i x_i \leq f_i \quad \forall i \tag{18}$$

and for the dual:

$$z = \sum_i p_i' y_i - \sum_i r_i' f_i \tag{19}$$

$$p_i'(I - A_i) - r_i' F_i \leq \partial_i' F_i \quad \forall i \tag{20}$$

$$p_i'(I - T_{ij}) - p_j \leq 0 \quad \forall i, j \in i \neq j \tag{21}$$

The basic static one-region model represented by Equations (1) – (4) can be used to calculate output, factor use, and goods prices under alternative assumptions about changes in the region's input structure for one or more sectors, its final deliveries, or its

factor prices. In the case of the world model with bilateral trade, the following variables can be calculated for each region: output, factor use, imports by origin region, and exports by destination region in the physical model, and region-specific goods prices and factor scarcity rents in the price model.

Now scenarios can accommodate assumptions not only about production techniques, factor prices, and consumption but also about factor endowments and characteristics of the global transport industries. Any number of resources that are typically ignored in economic models are readily introduced, for example arable land or fresh water, along with region-specific quantity constraints on their availability. Emissions matrices can also be applied to each regional output vector, with or without associated fees or other monetary costs.

4. Concluding Observations

 This paper has presented the basic input-output model and illustrated how the broad applicability and flexibility of the fundamental input-output concepts has resulted in a continuous evolution of the model to provide increasing explanatory power in diverse applications. It is always a challenge for a researcher to interpret or reinterpret a familiar model in such a way that it can provide new insight into precisely the question at hand. One way of doing that in the context of the economy is to incorporate into the model a representation of specific actors, sectors or activities that are known to interact with those already included. Sometimes the selection is dictated by the desire to address new kinds of applied questions, such as those related to energy use or environmental problems.

In such cases, the researcher will often attempt to maintain the existing model structure until reaching a limit where the model must be adapted. At this point difficult decisions may have to be made that involve rethinking system-wide relations between existing and new variables and defining the parameters that quantify these relationships. Other decisions involve the need to relax assumptions, for example by endogenizing relationships previously taken as exogenous or fixed.

The modeler is always confronted by the choice of parameters. In research models, they should be based in economic theory rather than just being the quotient of one variable divided by another. When they are based on economic theory, the parameters can be expected to change only slowly and in ways that are well understood conceptually.

It is of course possible to enlarge a model without changing its conceptual structure. Thus, for example, one could implement an input-output model with thousands of sectors or a world model with all of the approximately two hundred countries represented as potential trade partners. While the advantages of additional detail and disaggregation are evident, it should be recognized that there are also drawbacks. The relationships represented by the model's equations will often be more volatile; for example, if similar products are finely distinguished, the prospects for their substitutability cannot be ignored. Clearly the availability of adequate data becomes much more problematic, for example because of confidentiality.

These dilemmas facing the researcher regularly reappear in new guises. For example, some of the kinds of quandaries faced in regional input-output analysis now are now being faced and addressed in the new generation of world models. The researcher addressing new questions is continually engaged in a balancing act requiring judicious choices among the many options available. This is certainly the case for input-output economists.

BIBLIOGRAPHY

Brody, A. 1970. Proportions, Prices and Planning, North-Holland, Amsterdam. [Seminal text on input-output models of economic growth].

Chenery, H.B. 1953. "Regional Analysis." In H.B. Chenery, P.G. Clark, and V. Cao Pinna, eds., The Structure and Growth of the Italian Economy, Rome: U.S. Mutual Security Agency. [Introduces fundamentals of the MRIO model].

Duchin, F. 2005. "A World Trade Model based on Comparative Advantage with m Regions, n Goods, and k Factors," Economic Systems Research, 17(2): 141-162. [Trade and prices are based on comparative advantage with factor constraints.]

Duchin, F. 1988. "Analyzing Structural Change in the Economy." In M. Ciaschini, ed., Input-Output Analysis: Current Developments, London: Chapman & Hall. [Close of the basic static model is described for most components of final demand.]

Duchin, F. and G. Lange. 1995. "The Choice of Technology and Associated Changes in Prices in the U.S. Economy," Structural Change and Economic Dynamics, Vol. 6, No. 3: 335-357. [This is a one-region linear program for the United States the analysis of the choice among alternative technologies.]

Duchin, F. and D. Szyld. 1985. "A Dynamic Input-Output Model with Assured Positive Output," Metroeconomica, 37: 269-282. Reprinted in H. Kurz, C. Lager, and E. Dietzenbacher, eds., 1998, Input-Output Analysis, (Cheltenham: Edward Elgar Publishing Ltd). [The dynamic input-output model with capacity utilization.]

Hubacek, K. and L. Sun. 2001. "A Scenario Analysis of China's Land Use Change: Incorporating Biophysical Information into Input-Output Modeling," Structural Change and Economic Dynamics, 12(4): 367-397. [An example of a mixed-unit environmental analysis.]

Isard, W. 1951. "Interregional and Regional Input-Output Analysis: A Model of a Space Economy," Review of Economics and Statistics, 33(4): 318-328. [A fundamental contribution to regional analysis and in particular the IRIO model].

Jorgenson, D.W. 1998. "A Dual Stability Theorem," In D.W. Jorgenson, Growth, Econometric General Equilibrium Modeling, Vol. I, 1-8.: Cambridge, Mass., MIT Press. [Discusses stability problems of the dynamic Leontief models].

Jorgenson, D.W. 1998. "The Structure of Multi-Sector Dynamic Models," In D.W. Jorgenson, Growth, Econometric General Equilibrium Modeling, Vol. I, 41-60.: Cambridge, Mass., MIT Press. [Discusses multi-sector generalizations of Keynesian type models].

Julia, R. and F. Duchin. 2007. "World Trade as the Adjustment Mechanism of Agriculture to Climate Change," Climatic Change. [An example of a mixed-unit environmental analysis in a world trade model.]

Konijn, P.J.A. 1994. The Make and Use of Commodities by Industries; On the Compilation of Input-Output Data from the National Accounts. University of Twente, Enschede. [Provides a link between National Accounting Systems and input-output modeling].

Leontief, W.W. 1970. "Environmental repercussions and the economic structure: an input-output approach," Review of Economics and Statistics, 52, 262-271. [Introduces environmental pollution and abatement activities into the input-output structure].

Leontief, W. 1970. "The Dynamic Inverse." In A.P. Carter, and A. Brody, eds., Contributions to Input-Output Analysis: Proceedings of the Fourth International Conference on Input-Output Techniques, Vol. I: 17-46.: North-Holland Publishing Company. [The dynamic inverse matrix is presented.]

Leontief, W. 1937. "Interrelation of Prices, Output, Savings and Investment: A Study in Empirical Application of Economic Theory of General Interdependence." Review of Economic Statistics, 19(3):109-32. [One of the pair of articles introducing input-output economics.]

Leontief, W. 1936. "Quantitative Input and Output Relations in the Economic System of the United States." Review of Economics and Statistics, 18(3):105-25. [One of the pair of articles introducing input-output economics.]

Leontief, W.W. and D. Ford. 1972. "Air pollution and the economic structure: empirical results of input-output computations." In A. Brody and A.P. Carter, eds., Input-Output Techniques, Amsterdam: North-Holland, 9-30. [Presents calculations on the basis of the Leontief environmental model].

Miller, R.E. and P.D. Blair. 1985. Input-Output Analysis: Foundations and Extensions, Englewood Cliffs, Prentice-Hall. [A basic text on input-output modeling].

Moses, L.N. 1955. "The Stability of Interregional Trading Patterns and Input-Output Analysis," American Economic Review 45(5): 803-832. [One of the fundamental texts presenting the MRIO model].

Polenske, K.R. 1980. The U.S. Multiregional Input-Output Accounts and Model, Lexington, Mass., Lexington Books. [A detailed application of the U.S. MRIO model].

Polenske, K.R. and G. Hewings. 2004. "Trade and Spatial Economic Interdependence," Papers in Regional Science, 83: 269-289. [Links theoretical and applied regional modeling frameworks with a focus on trade and development].

Rose, A. and S.D. Casler. 1996. "Input-Output Structural Decomposition Analysis: A Critical Appraisal," Economic Systems Research, 8, 33-62. [An appraisal of decomposition techniques].

Steenge, A.E. 1978. "Environmental repercussions and the economic structure: further comments," Review of Economics and Statistics, 60, 482-486. [Discusses pricing problems in the Leontief environmental model].

Steenge, A.E. 2005. "Social cost in the Leontief environmental model: rules and limits to policy," In E. Dietzenbacher and Michael Lahr, eds., Wassily Leontief and Input-Output economics, Cambridge: Cambridge University Press. [Discusses the limits to what environmental policy can do].

Steenge, A.E. and M.J.P.M. Thissen. 2005. "A New Matrix Theorem: Interpretation in Terms of Internal Trade Structure and Implications for Dynamic Systems." Journal of Economics, 84, 71-94. [Provides an alternative approach to stability problems of the dynamic input-output models].

Strømman, A. and F. Duchin. 2006. "A World Trade Model with Bilateral Trade Based on Comparative Advantage." Economic Systems Research, 18(3), September. [The world trade model with endogenous transportation and bilateral trade flows.]

Strømman, A. E. Hertwich, and F. Duchin. In review. "Shifting Trade Patterns as a Means of Reducing Global CO2 Emissions: From A Global Objective to Industry-Specific Implications." [Analysis with economic and environmental objective functions and a trade-off curve of cost per increasing increment of carbon reduction.]