

# A non-linear Leontief–type input-output model

Panayotis G. Michaelides<sup>\*</sup>, Athena Belegri-Roboli *and* Maria Markaki

*National Technical University of Athens*

**ABSTRACT** – In this paper, we present an econometric model based on Leontief’s Input–Output (IO) approach. In this context, the paper puts forward an econometric approach to estimating an IO Leontief – type coefficients matrix which has several advantages and constitutes an extension to the standard IO model. Analytically, the original model is a description of the situation when (i) linear relations express the production process of each sector, (ii) each sector experiences constant returns to scale, and (iii) the technical coefficients in the conventional IO table are fixed for several years and based on *a-priori* calculations using traditional survey methods made by practitioners, and not on econometric estimations using real–world data on economic aggregates. The proposed method’s main advantage is its simplicity, flexibility, and capability of including real-world information on economic aggregates that could also be used as a portion of a dynamic model. Measures such as Returns to Scale (RTS), Total Factor Productivity (TFP), and Technical Efficiency (TE) may be computed easily.

*Keywords:* Leontief, input-output model, econometrics, RTS, TFP, TE

## I. Introduction

There is no doubt that “the biggest challenge for researchers remains the issue of the appropriate production function specification to represent the underlying process technology” (Vaneman and Triantis, 2007). In this context, input – output analysis has been widely used as an analytical tool since its development (Sharp and Perkins, 1973). In fact, the Leontief input-output model is probably the “most well-known and most often used static model of the structure of a national economy” (Chander 1983: 219). The

---

<sup>\*</sup> P. G. Michaelides, Ass. Professor of *Economics*, Laboratory of Theoretical & Applied Economics, School of Applied Mathematics & Physics, National Technical University of Athens, Greece, Heron Polytechniou 9, 157.80, Zografou Campus, Tel: 302107721624; fax: +302107721618, email: pmichael@central.ntua.gr (*Contact Author*).

input – output (IO) model originally introduced by Leontief (1941) provides a linear characterization of the production technology of any given economic system, within a multisector economy. It, thus, constitutes a key approach in economics and management.

However, many other functional forms have also been proposed for the approximation of a production technology. Perhaps the most popular in current use is the Cobb-Douglas (CD) specification introduced by Cobb and Douglas (1928) that is intuitively appealing and computationally straightforward, and has been extensively used to test various restrictions of production theory. In this context, one has to balance between the costs of using a conventional CD specification which does not embody inter-sectoral linkages and the apparent benefits, such as increased flexibility, better approximations and predictions, more accurate decision making, etc.

It should be made clear that in the context of modeling a production process, the non-linear CD specification can be used to obtain a sufficiently flexible functional form which is capable of approximating an existing production process of a multisector economy. Meanwhile, the multisectoral nature of the IO model makes it quite attractive in production theory; however it is very restrictive in terms of flexibility. Needless to say, it would be extremely helpful to estimate a functional form that combines, in some sense, both approaches appropriate for modeling multi-sectoral production technologies with a sufficiently flexible functional form.

The conventional IO model is a description of the situation when (i) linear relations express the production process of each sector, (ii) each sector experiences constant returns to scale, and (iii) the technical coefficients in the IO table are fixed for several years and based on *a-priori* calculations using traditional survey methods made by practitioners, and not on econometric estimations based on real-world data on economic aggregates. The proposed method's main advantage is its simplicity, flexibility, and capability of including real-world information on economic aggregates that could also be used as a portion of a dynamic model. Also, the problems that are associated with the construction of IO tables using

conventional survey methods make the development of a relevant technique very attractive. Moreover, IO tables are usually available with delay. As a result, estimates of the coefficients have to be used.

So far, several researchers have attempted to extend the linear IO model to account for non-linearity in the production process, in an effort to make it more realistic (e.g. Nikaido 1968, Sandberg 1973, Chander 1983, Tokutsu 1984, Fujimoto 1986, Read 1986, Zhang 2000, Zhao et al. 2006). However, probably the most serious drawback of these approaches is the fact that they try to introduce the stochastic element through direct estimation of the basic IO linear equation (see, in this spirit, Gerking 1976a, b, c, 1979a, b, Hanseman and Gustafson, 1980, 1981 among others) and *not* by introducing some other flexible specification that could be estimated simultaneously with the traditional IO equation. As a result, these contributions have found extremely limited applications in the literature. After all (see Kocklaeuner 1989), econometric analysis of input-output models which started with Gerking (1976) has practically “not entered input-output analysis”, with some exceptions (see e.g. Bulmer-Thomas, 1982).

Of course, no single model can properly capture *all* of the features of economic reality. However, by focusing on some of them, an appropriate model can be constructed and estimated by relevant techniques. In this paper, we propose a new model which has the following advantages: (i) It provides a good approximation to any arbitrary multi-sectored production process; (ii) it is based on functional specifications which are consistent with known production theory; (iii) it is flexible with respect to time; (iv) it allows for arbitrary returns to scale (RTS) and total factor productivity (TFP); (v) it allows for Technical Efficiency (TE) estimation, (vi) it is simple to estimate; and (vii) can be used as a portion of a dynamic model (Sandberg, 1973). We believe that the proposed model is an appropriate vehicle for expanding conventional Leontief’s IO analysis.

The paper is organised as follows: Section II introduces the model. Section III provides the framework for measuring RTS, TFP and TE. The procedure for the empirical estimation of the model is set out in Section IV, while Section V concludes.

## II. The Model

### A. Theoretical formulation

The conventional IO model is based on the following equation for the various economic sectors:

$$X_i = x_{i1} + x_{i2} + \dots + x_{in} + y_i, \quad i, j = 1, 2, \dots, n \quad (1)$$

where:  $X_i \geq 0$  is the output of sector  $i$ ,  $y_i$  is the final demand for the product of sector  $i$ ,  $x_{ij}$  is the product of sector  $i$  used by sector  $j$ .

Equation (1) can be written in matrix form as follows:

$$X = AX + Y \quad (2)$$

where:  $X$  is the vector of outputs,  $Y$  is the vector of final demand, and  $A$  is the so-called technical coefficients matrix whose typical element is equal to :

$$(\alpha_{ij})_{n \times n} = \frac{x_{ij}}{X_j} \quad (3)$$

where:  $(\alpha_{ij})_{n \times n} \geq 0$  is interpreted as the quantity of output from sector  $i$  required to produce one unit of output in sector  $j$

As we know, in the conventional IO model the main tool of analysis is the so-called Leontief inverse  $(I - A)^{-1}$ , namely the matrix of input-output multipliers of changes in final demand into levels of outputs.

However, this approach has several drawbacks. A first major criticism lies in the fact that the production process in any economic system in general is non-linear, while Leontief's conventional IO model is linear. In this sense, a sufficiently flexible non-linear production specification is needed. Second, an implicit assumption in the typical IO model leads to the conclusion that the technical coefficients in the IO table are fixed and based on *a-priori* calculations using traditional survey methods made by practitioners,

and not on econometric estimations based on real–world data on economic aggregates. However, as we know, with the development of technology, the changes in these coefficients are significant and have to be estimated based on real–world data over a period of time. Third, the original model is a description of the situation when each sector of economic activity experiences constant RTS. However, this assumption is not expected to approximate reality with any reasonable accuracy and is always an empirical matter. In addition, interesting possibilities which arise when there are non-constant RTS, are not amenable to the Leontief input-output model. In this context, a non-linear production function specification has to be used and estimated, based on a combination of conventional IO theory and production function theory resulting in a new Leontief–type model (and table), respectively, which has considerable advantages.

Motivated by the considerations outlined earlier, in this paper the main idea is to express the output  $X_i$  of any given sector  $i$  as a function of:

(a) the conventional input factors such as physical capital  $K_i$ , labor  $L_i$  and human capital  $H_i$  which are used in the production process of sector  $i$ , and

(b) other sectors' output which is used as an input in the production process of sector  $i$ .

In a general formulation:

$$X_i = f(K_i, L_i, H_i, X_1, X_2, \dots, X_j), \quad j = 1, 2, \dots, n, \quad j \neq i$$

The  $f$  function can be easily specified as a CD production function. So, we assume that the production process of any given sector can be characterised by the non-linear flexible CD specification which is intuitively appealing and computationally straightforward and has been estimated extensively.

In this context, the proposed production function is given by the following specification:

$$X_i = N_i K_i^{k_i} L_i^{l_i} H_i^{h_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t} \quad (4)$$

where:  $t$  is a time index and  $N_i \geq 0$ ,  $k_i \geq 0$ ,  $l_i \geq 0$ ,  $h_i \geq 0$ ,  $c_{ji} \geq 0$ ,  $i, j = 1, 2, \dots, n, j \neq i$  are parameters.

In an IO spirit, the part of the output of any given sector  $j$  that is used as an input in the production process of some other sector's output  $i$ , namely  $X_j^{c_{ji}}$  is practically equal to  $x_{ji}$ .

In other words:

$$x_{ji} = X_j^{c_{ji}}, i, j = 1, 2, \dots, n, j \neq i \quad (5a)$$

Thus:

$$x_{ii} = X_i^{c_{ii}}, i, j = 1, 2, \dots, n, j = i \quad (5b)$$

Apparently:

$$x_{ji} \neq x_{ij}, j \neq i, i, j = 1, 2, \dots, n$$

$$x_{ii} \neq x_{jj}, j \neq i, i, j = 1, 2, \dots, n$$

So, the proposed non linear Leontief-type model which describes the economic system, and is essentially different from the conventional one, consists in the following equations:

$$X_i = x_{i1} + x_{i2} + \dots + x_{in} + y_i, i, j = 1, 2, \dots, n \quad (1)$$

$$X_i = N_i K_i^{k_i} L_i^{l_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t}, i, j = 1, 2, \dots, n, j \neq i \quad (4)$$

which is equivalent to the following non linear Leontief-type model system:

$$X = AX + Y \quad (2)$$

$$(\alpha_{ij})_{n \times n} = \frac{x_{ij}}{X_j} \quad (3)$$

$$X_i = N_i K_i^{k_i} L_i^{l_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t}, j \neq i \quad (4)$$

The second equivalent formulation is consistent with the IO model and is adopted because of its convenience to calculate the Leontief-type  $A$  matrix. Thus, after estimating the parameters of interest via equations (1) & (4), the  $A$  matrix is calculated easily via equation (3) and is applied to equations (2).

### III. Measure of Interest

#### A. Returns to Scale

As we know, returns to scale (RTS) refer to a technical property of the production process that examines changes in output subsequent to a proportional change in the own inputs of each sector. If output increases by the same proportional change then there are constant returns to scale (CRS). If output increases by less than that proportional change, there are decreasing returns to scale (DRS). If output increases by more than that proportion, there are increasing returns to scale (IRS). In other words, if  $RTS < 1$  ( $> 1$ ) the production technology is characterised by DRS (IRS). If  $RTS = 1$  we have CRS.

The proposed approach does not place *a priori* restrictions on the behaviour of RTS like other functional forms. As we know, the RTS are equal to the sum of the output elasticities of the sector's own inputs. Based on Equation (4), the RTS are equal to:

$$RTS_i = k_i + l_i, j = 1, 2, \dots, n \text{ and } j \neq i \quad (6)$$

#### B. Total Factor Productivity

Growth in TFP represents output growth not accounted for by the growth in inputs and explains changes in productivity, i.e. the production of more output with a given level of inputs (Varian, 1992). TFP, by definition is equal to:

$$TFP_i = \frac{\partial \ln X_i(x)}{\partial t}$$

and based on Equation (4) we have that :

$$TFP_i = \lambda_i \quad (7)$$

### C Technical Efficiency Estimation

As we know, technical efficiency analysis is an indispensable tool for evaluating an economic unit's performance. Consequently, reliable measures of technical efficiency are of great interest. If we follow the stochastic frontier approach proposed by Aigner *et al.* (1977) and Meeusen and van der Broeck (1977) we should add a symmetric error term  $e$  to capture the effects of white noise:

$$\ln X_i = \ln N_i + k_i \ln K_i + l_i \ln L_i + h_i \ln H_i + c_{1i} \ln X_1 + c_{2i} \ln X_2 + \dots + c_{ji} \ln X_j + \lambda_i t + \varepsilon_i + u_i, j=1, 2, \dots, n, j \neq i \quad (8)$$

which is the expression of a stochastic frontier model including  $\varepsilon_i$  to capture the effects of white noise and  $u_i$  is a non-negative term such that  $0 < u_i \leq 1$ ,  $-\infty < \ln u_i \leq 0$  that captures the effects of inefficiency (Kumbhakar and Lovell, 2000).

As we know, the estimation of TE is based on the idea that the data are contaminated with measurement errors and other noise (e.g. Bauer, 1990). In brief, the well-known idea behind this specification is that production is subject to two disturbances. Namely, that the positive disturbance  $u_i$  expresses the fact that each firm's output lies on or below its frontier, whereas any deviation is regarded as the result of factors controllable by the firm. Meanwhile, the frontier itself may vary randomly over time for the same firm with random disturbance  $\varepsilon_i$ , which expresses events beyond the control of the firm.

In this approach, the typical assumption about equation (8) is that  $\varepsilon$  are *iid*  $(0, \sigma^2)$  and uncorrelated with the regressors. No distributional assumption is needed on the  $u_i$ . It is only required that  $u \geq 0$ , which represents inefficiency (Kumbhakar and Lovell, 2000). However, no observation may lie above the frontier. So, we have to ensure that the estimated frontier bounds the data from above and we have to subtract the largest residual from each residual in the sample (Kumbhakar and Lovell, 2000):

$$-\hat{u}_i^* = \hat{u}_i - \max_i \{\hat{u}_i\} \quad (9)$$

In this typical approach of measuring efficiency, the residuals  $\hat{u}_i^*$  are non-negative, with at least one being zero, and are used to provide consistent estimates of technical efficiency of each firm. This implies that  $-\hat{u}_i^* \leq 0$  and, therefore, they can be used as inefficiency measures (Kumbhakar and Lovell, 2000). This analysis is applied to the residuals of Equations (8). In other words, the estimate of the intercept of equation (8) is adjusted, so that the function no longer passes through the centre of the observed points but bounds them from above and the distance measure for the  $i$ th economic unit is then calculated as the exponent of the (corrected) residual. Thus, for any given year,  $TE$  is equal to:

$$TE_i = \exp(-\hat{u}_i^*) \quad (10)$$

This model is attractive because of its simplicity. Technical efficiency (TE) could be estimated in various ways, where distributional assumptions on the two error components have to be made. For instance, the conventional assumption is that  $u \sim N^+(0, \sigma^2)$  which is typically employed in empirical works. However, other distributional assumptions on the one-sided error component are employed (e.g. exponential, normal, truncated normal, gamma, etc), but less frequently because of their increased computational complexity (Kumbhakar and Lovell, 2000).

Of course, the average annual value of  $TE_i = \exp(-\hat{u}_i^*)$  by sector gives the arithmetic average of the technical efficiency measure that prevailed in sector  $i$  over the period under examination.

## IV. Empirical Analysis

### A. Econometric Estimation

As we have seen, the proposed non linear Leontief-type model consists in the following equations:

$$X_i = N_i K_i^{k_i} L_i^{l_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t}, \quad i, j = 1, 2, \dots, n, j \neq i \quad (4)$$

$$X_i = x_{i1} + x_{i2} + \dots + x_{in} + y_i, \quad i, j = 1, 2, \dots, n \quad (1)$$

where:

$$x_{ji} = X_j^{c_{ji}}, i, j = 1, 2, \dots, n, j \neq i \quad (5a)$$

$$x_{ii} = X_i^{c_{ii}}, i, j = 1, 2, \dots, n, j = i \quad (5b)$$

which is equivalent to the following non linear Leontief-type model system:

$$X = AX + Y \quad (2)$$

$$(\alpha_{ij})_{n \times n} = \frac{x_{ij}}{X_j} \quad (3)$$

$$X_i = N_i K_i^{k_i} L_i^{l_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t}, j \neq i \quad (4)$$

Now, equations (1), based on equations (5), yield equations (11):

$$X_i = X_1^{c_{i1}} + X_2^{c_{i2}} + \dots + X_n^{c_{in}} + y_i, i, j = 1, 2, \dots, n \quad (11)$$

Thus, the final system of equations to be estimated is the following:

$$X_i = N_i K_i^{k_i} L_i^{l_i} X_1^{c_{1i}} X_2^{c_{2i}} \dots X_j^{c_{ji}} e^{\lambda_i t}, i, j = 1, 2, \dots, n, j \neq i \quad (4)$$

$$X_i = X_1^{c_{i1}} + X_2^{c_{i2}} + \dots + X_n^{c_{in}} + y_i, i, j = 1, 2, \dots, n \quad (11)$$

Equations (4) express the production functions and the Equations (11) practically express cross-equations restrictions. In this context, the next step is to convert Equations (4) into an estimable model. In this context, an alternative form of (4) is:

$$\ln X_i = \ln N_i + k_i \ln K_i + l_i \ln L_i + h_i \ln H_i + c_{1i} \ln X_1 + c_{2i} \ln X_2 + \dots + c_{ji} \ln X_j + \lambda_i t, j = 1, 2, \dots, n, j \neq i \quad (12)$$

At this point, we can use regression analysis on a sample of data in order to econometrically estimate the unknown parameters. Equations (4) could be estimated by various techniques, including, of course, single – equation estimation based, for instance, on Ordinary Least Squares (O.L.S.) applied to Equations (12). After all, the O.L.S. estimators have considerable advantages. Namely, they have statistical properties at least as desirable (Olson *et al.* 1980) as those of the Maximum Likelihood (M.L.) estimators (Aigner *et al.*

1977), are easier to obtain and tend to provide encouraging results (Kumbhakar and Lovell, 2000). Also, this method is attractive due to its simplicity, which is extremely crucial when a large number of equations are considered that increase the computational load. Moreover, this method has the considerable advantage that if there is some degree of misspecification it does not spread to all the equations and it is contained within the equation that is misspecified.

Alternatively, we can estimate Equations (1) and (4) as a system of simultaneous equations. In this framework, the three Stages Least Square (3SLS) technique can be employed. The system is:

$$\left\{ \begin{array}{l} \ln X_1 = \ln N_1 + l_1 \ln L_1 + k_1 \ln K_1 + h_1 \ln H_1 + c_{21} \ln X_2 + \dots + c_{n1} \ln X_n + \lambda_1 t \\ \dots \\ \ln X_n = \ln N_n + l_n \ln L_n + k_n \ln K_n + h_n \ln H_n + c_{1n} \ln X_1 + \dots + c_{nn} \ln X_n + \lambda_2 t \\ X_1 = X_1^{c_{11}} + X_1^{c_{12}} \dots + X_1^{c_{1n}} + y_1 \\ \dots \\ X_n = X_n^{c_{n1}} + X_n^{c_{n2}} + \dots + X_n^{c_{nn}} + y_n \end{array} \right. \quad (13)$$

The 3SLS technique obtains three stage least squares estimates of a set of linear or nonlinear equations. 3SLS is at least as efficient as any other estimator which uses the same amount of information and it is also suitable when there might be error correlation in different equations. 3SLS estimates are consistent and asymptotically normal, and, can be asymptotically more efficient than single equation estimates. Technically, 3SLS is a combination of Seemingly Unrelated Regression (SUR) estimation and two Stages Least Square (2SLS). It takes into account the covariances across equation disturbances. The objective function for 3SLS is the sum of squared transformed fitted residuals (e.g. Davidson and MacKinnon, 1993; Theil, 1973).

From a methodological point of view, 3SLS are obtained by estimating a set of (non-)linear equations with cross-equation constraints imposed, but with a diagonal covariance matrix of the disturbances across equations. The parameter estimates thus obtained are used to form a consistent estimate of the covariance matrix of the disturbances, which is then used as a weighting matrix when the model is re-

estimated to obtain new values of the parameters. The method of parameter estimation is the so-called iterative Gauss-Newton method for nonlinear least squares (NLS), which is quite well-known in fields such as Numerical Analysis, etc. For details on the properties of the linear 3SLS estimator see the seminal work by Zellner and Theil (1962). For the nonlinear 3SLS estimator, see the benchmark works by Amemiya (1977) and Jorgenson and Laffont (1975).

## V. Conclusion

In this paper, a non-linear stochastic IO model is put forward which constitutes an extension of the conventional Leontief model. The original model is a description of the situation when: (i) linear relations govern the production process of each sector, (ii) each sector experiences CRS and (iii) the technical coefficients in the IO table are fixed for several years and based on *a-priori* calculations using conventional survey methods made by practitioners, and not on econometric estimations based on real-world data on economic aggregates, over a period of time. Of course, in real-life organizational and economic systems, things are much more complicated and these assumptions are not expected to express economic reality with any reasonable accuracy. Also, the problems that are associated with the construction of IO tables using conventional survey methods make the development of a relevant technique very attractive.

In this context, we propose a new production function equation which, when combined with the conventional IO equations, provides a more complete, accurate and updated description of the economy. Of course, the advantages of Leontief's original IO model have been fully retained. Also, we all measures such as RTS, TFP, and TE are computed easily. Finally, based on our estimates an IO table based on econometric estimations of real-world data on economic aggregates, over a period of time, could be estimated.

To conclude, in this paper we proposed a new model which has considerable advantages: (i) It gives an approximation to any arbitrary multi-sectoral production process; (ii) it is flexible with respect to time; (iii) it allows for arbitrary Returns To Scale (RTS) and Total Factor Productivity (TFP); (iv) it is

simple to estimate; (v) it has a functional form which is consistent with known production theory; (vi) it allows for Technical Efficiency (TE) estimation. Finally, (vii) such models can be used as portions of dynamic models. Although some of these desirable properties are also possessed by one or other of the known approaches, neither possesses all of them simultaneously.

Of course, there are several issues that could serve as good examples for future investigation. For instance, an application of the proposed approach poses an interesting challenge. Also, one could explore other production function specifications such as the Translog, etc. Moreover, one could try to incorporate other input factors and variables such as Human Capital, Research and Development, etc. In a theoretical framework, one could investigate the properties of the Leontief-type matrix. No doubt, future and more extended research on the subject would be of great interest.

## References

Amemiya, Takeshi, "The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equation Model," *Econometrica*, May 1977, pp. 955-975.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman, "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, October 1975, pp. 653-665.

Davidson, R. and J. G. MacKinnon, *Estimation and Inference in Econometrics*, Oxford University Press, 1993.

Gallant, A. Ronald, and Dale W. Jorgenson, "Statistical Inference for a System of Simultaneous, Non-linear, Implicit Equations in the Context of Instrumental Variable Estimation," *Journal of Econometrics* 11, 1979, pp. 275-302.

Jorgenson, Dale W. and Jean-Jacques Laffont, "Efficient Estimation of Nonlinear Simultaneous Equations with Additive Disturbances," *Annals of Economic and Social Measurement*, October 1975, pp. 615-640.

Theil, Henri, *Principles of Econometrics*, John Wiley and Sons, New York, 1971, pp. 508-527.

Zellner, Arnold, and Henri Theil, "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica* 30 (1962), pp. 54-78.