CHAPTER 4

THEORETICAL FRAMEWORK

The relationship between productivity and economic growth has been debated extensively in the literature. In a neoclassical framework, economic growth stems from either factor accumulation or growth in Total Factor Productivity (TFP). Although initially the most common method of calculating Total Factor Productivity Growth (TFPG) was growth accounting, both index number and distance function methods for calculating change in TFP levels are becoming more common. This chapter discusses the theoretical framework for growth accounting, for productivity indexes and also for the distance function approach of productivity measurement.

4.1 Measuring Total Factor Productivity Growth using Growth Accounting

4.1.1 Total Factor Productivity Growth

Solow (1957) defined total factor productivity growth as the rate of growth of real output not accounted for by the growth of the factor inputs and associated it with a shift in technology. When considered at some particular point of time *t*, an aggregate production function links output Y_t to the factors of production like the stock of capital K_t and the size of the labor force N_t and also to the total factor productivity A_t. The Cobb-Douglas form of such a production function would be

$$Y_t = A_t \,.\, K_t^{\alpha} \,.\, N_t^{1-\alpha} \tag{4.1}$$

Here, the change in the output is possible only through the change in the capital stock, the size of the labor force, or the level of TFP. In growth accounting, therefore, TFP is calculated as a geometric index:

$$A_t = Y_t \,/\, (K_t^{\alpha} \,.\, N_t^{1-\alpha}) \tag{4.2}$$

Taking natural log on both sides of equation 4.2 will result in equation 4.3.

$$ln\ A_t = ln\ Y_t - \alpha\ .\ ln\ K_t\ - (1\text{-}\alpha)\ .\ ln\ N_t \qquad (4.3)$$

Consequently, the TFPG is the difference between the rate of output growth and the contribution from the input factors. There are three views about interpreting this TFPG (Carlaw and Lipsey, 2003):

(a) TFPG is 'advance of knowledge' or 'technical change' or 'technical progress' as improvement in technology would increase productivity.
(b) TFPG is 'free lunches' being external to the economic activities that generate growth.
(c) TFPG is the 'measure of our ignorance', since it represents the portion of the increase in output that cannot be accounted for by the increase in inputs considered in the analysis.

However, TFPG is often equated with technical change in the literature.

4.1.2 Growth Accounting [1]

The analysis of growth accounting is based on the concept of an aggregate production function exhibiting the relationship between output and input factors. A general form of production function can be written as equation 4.4.

$$Y(t) = f\,(x(t),\ t) \qquad (4.4)$$

where *Y(t)* is output at time *t* and *x(t)* is input at time *t*.

The time variable *t* is included in the function to represent the possible shift of the production function over time. Assuming that there are two inputs namely capital expressed as *K* and labor expressed as *N*, the equation 4.4 can be rewritten as equation 4.5.

[1] The illustration for growth accounting shown in this section and afterwards is taken from Tinakorn and Sussangkarn (1998).

$$Y(t) = f\,(K(t), N(t), t) \tag{4.5}$$

This production function implies that the increase in output Y must be attributable to the increase in inputs and/or the shift in the production technique. Now, taking total differentiation of Y with respect to time t in equation 4.5,

$$\frac{dY}{dt} = \frac{\partial f(.)}{\partial K} \cdot \frac{dK}{dt} + \frac{\partial f(.)}{\partial N} \cdot \frac{dN}{dt} + \frac{d\,f(.)}{dt} \tag{4.6}$$

Equation 4.6 can be written as,

$$\dot{Y} = \frac{\partial f(.)}{\partial K} \cdot \dot{K} + \frac{\partial f(.)}{\partial N} \cdot \dot{N} + \dot{f}(.) \tag{4.7}$$

Here, the dot over a variable indicates differentiation with respect to time. Dividing equation 4.7 by equation 4.5 and rearranging terms will obtain equation 4.8 and equation 4.9 respectively.

$$\frac{\dot{Y}}{Y} = \frac{\partial f(.)}{\partial K} \cdot \frac{\dot{K}}{f(.)} + \frac{\partial f(.)}{\partial N} \cdot \frac{\dot{N}}{f(.)} + \frac{\dot{f}(.)}{f(.)} \tag{4.8}$$

$$\frac{\dot{Y}}{Y} = \frac{\partial f(.)}{\partial K} \cdot \frac{K}{f(.)} \cdot \frac{\dot{K}}{K} + \frac{\partial f(.)}{\partial N} \cdot \frac{N}{f(.)} \cdot \frac{\dot{N}}{N} + \frac{\dot{f}(.)}{f(.)} \tag{4.9}$$

But it is notable that, in the equation 4.9,

$$\frac{\partial f(.)}{\partial K} \cdot \frac{K}{f(.)} = \frac{\partial f(.)\,/\,f(.)}{\partial K\,/\,K}$$ = output elasticity with respect to capital (β_K)

$$\frac{\partial f(.)}{\partial N} \cdot \frac{N}{f(.)} = \frac{\partial f(.)\,/\,f(.)}{\partial N\,/\,N}$$ = output elasticity with respect to labor (β_N)

So, equation 4.9 can be written as equation 4.10.

$$\frac{\dot{Y}}{Y} = \beta_K \cdot \frac{\dot{K}}{K} + \beta_N \cdot \frac{\dot{N}}{N} + \frac{\dot{f}(.)}{f(.)} \tag{4.10}$$

Equation 4.10 states that the rate of change in output ($\dot{Y}/Y$) can be decomposed into two major parts:

(i) the sum of the rate of change in inputs weighted by their respective output elasticity expressed as

$$\beta_K \cdot \frac{\dot{K}}{K} + \beta_N \cdot \frac{\dot{N}}{N}$$

and

(ii) the shift of production function through time expressed as

$$\dot{f}(.) / f(.)$$

The latter part can be written as the difference between the rate of output growth and the contribution from inputs following equation 4.10.

$$\frac{\dot{f}(.)}{f(.)} = \frac{\dot{Y}}{Y} - \beta_K \cdot \frac{\dot{K}}{K} - \beta_N \cdot \frac{\dot{N}}{N} \qquad (4.11)$$

But the right hand side of equation 4.11 is the total factor productivity growth, so that

$$\text{TFPG} = \frac{\dot{Y}}{Y} - \beta_K \cdot \frac{\dot{K}}{K} - \beta_N \cdot \frac{\dot{N}}{N} \qquad (4.12)$$

From this point, the analysis can undertake any of the three approaches:

(i) The analysis can adopt a parametric approach, where values of β_K and β_N will be obtained from estimation of any particular form of production function like Cobb-Douglas function and the estimated values will be used in equation 4.12 to calculate the TFPG.

(ii) The analysis can adopt a non-parametric approach, where the assumptions of profit maximization and producer's equilibrium will allow the use of income shares going to capital and labor to replace values of β_K and β_N respectively, so that these replaced values can be used in equation 4.12 to calculate the TFPG.

These approaches are discussed separately below.

4.1.3 Parametric Approach to Growth Accounting

In the parametric approach to growth accounting, the output elasticity with respect to each input (β_K and β_N) is obtained from the estimation of any particular form of production function. For an illustration, taking natural logarithm and total differentiating equation 4.5 with respect to *t* would result in an expression,

$$\frac{d\,lnY}{dt} = \frac{\partial\,lnY}{\partial\,lnK} \cdot \frac{\partial\,lnK}{\partial t} + \frac{\partial\,lnY}{\partial\,lnN} \cdot \frac{\partial\,lnN}{\partial t} + \frac{\partial lnY}{\partial t} \tag{4.13}$$

$$\frac{dY}{dt} \cdot \frac{1}{Y} = \frac{\partial\,lnY}{\partial\,lnK} \cdot \frac{1}{K} \cdot \frac{\partial K}{\partial t} + \frac{\partial\,lnY}{\partial lnN} \cdot \frac{1}{N} \cdot \frac{\partial N}{\partial t} + \frac{1}{Y} \cdot \frac{\partial Y}{\partial t} \tag{4.14}$$

$$\frac{\dot{Y}}{Y} = \beta_K \cdot \frac{\dot{K}}{K} + \beta_N \cdot \frac{\dot{N}}{N} + TFPG \tag{4.15}$$

This is the same expression as equation 4.10, and β_K and β_N will have the expression as below,

$\beta_K = \partial\, lnY / \partial\, lnK$ = output elasticity of capital

$\beta_N = \partial\, lnY / \partial\, lnN$ = output elasticity of labor

Now, assuming a Cobb-Douglas production function of the form and taking natural log on both sides,

$$Y = A \,.\, K^{\beta K} \cdot N^{\beta N} \tag{4.16}$$

$$lnY = lnA + \beta_K \,.\, lnK + \beta_N \cdot lnN \tag{4.17}$$

With assumption of constant returns to scale,

$$\beta_K + \beta_N = 1 \tag{4.18}$$

So, equation 4.17 can be written as,

$$lnY = lnA + \beta_K \, . \, lnK + (1 - \beta_K) \, . \, lnN \tag{4.19}$$

$$lnY - lnN = lnA + \beta_K \, . \, (lnK - lnN) \tag{4.20}$$

This will enable one to estimate value of β_K from a time series data set using equation 4.20, to be used in equation 4.15 for calculating TFPG. However, the parametric approach requires a long enough series of data for obtaining a statistically reliable estimate of the parameters. Besides, incorrectly measured growth rates of input factors will lead to inconsistent estimates of factor income shares.

Another form of the parametric approach is direct regression, where the growth rate of output will be regressed on the growth rate of inputs in equation 4.10 so that the intercept will measure TFPG and the coefficients of the factor growth rates will measure β_K and β_N as output elasticity with respect to capital and labor respectively (Barro, 1999). The main advantage of this approach is that it dispenses with the assumption that the factor social marginal products coincide with the observable factor prices. A major disadvantage of this approach is that, the growth rate of input factors are regarded as exogenous with respect to growth of TFP because of which the factor growth rates would receive credit for correlated variations in unobservable technological change.

4.1.4 Non-Parametric Approach to Growth Accounting

In non-parametric approach, the problem of estimation of parameters is subdued by the assumptions of profit maximization and producers' equilibrium whereby each factor input is employed up to the point where its marginal product (MP) equals its real cost.

If W_i is price of input i and P is price of output, the marginal productivity of capital and labor would be as in equation 4.21 and equation 4.22 respectively.

$$MP_K = \partial f(.) / \partial K = W_K / P \tag{4.21}$$

$$MP_N = \partial f(.) / \partial N = W_N / P \tag{4.22}$$

Then with the assumptions of perfect competition in the market and constant returns to scale, the output elasticity of capital and the output elasticity of labor would be equal to the factor income share going to capital and labor respectively (s_K and s_N), as shown in equation 4.23 and 4.24 respectively.

$$\beta_K = \frac{\partial f(.)}{\partial K} \cdot \frac{K}{f(.)} = \frac{W_K}{P} \cdot \frac{K}{f(.)} = (W_K \cdot K) / (P.Y) = s_K \tag{4.23}$$

$$\beta_N = \frac{\partial f(.)}{\partial K} \cdot \frac{N}{f(.)} = \frac{W_N}{P} \cdot \frac{N}{f(.)} = (W_N \cdot N) / (P.Y) = s_N \tag{4.24}$$

In principle, the factor income shares can be obtained from the national income account. Hence, the task of calculating TFP growth at the national level appears rather straightforward as in equation 4.25.

$$\text{TFPG} = \frac{\dot{Y}}{Y} - s_K \cdot \frac{\dot{K}}{K} - s_N \cdot \frac{\dot{N}}{N} \tag{4.25}$$

where s_K is share of income going to capital and s_N is share of income going to labor.

Young (1992) refers to the TFP part of growth as the 'intensive' component of growth and the part from the contribution of factor inputs as the 'extensive' part. In an agriculture-based economy where land expansion is still possible, the land should be added as an additional primary input. In that case, the input vector can be written as,

$$x(t) = (K(t), N(t), L(t)) \tag{4.26}$$

Here, L stands for land. The same mathematical analysis as above will result in the expression,

$$\text{TFPG} = \frac{\dot{Y}}{Y} - s_K . \frac{\dot{K}}{K} - s_N . \frac{\dot{N}}{N} - s_L . \frac{\dot{L}}{L} \tag{4.27}$$

where s_L is rental share of income going to land owners.

Assumption of constant return to scale will imply that, sum of income shares will be one.

$$s_N + s_K + s_L = 1 \tag{4.28}$$

Since equations 4.25 and 4.27 are in continuous forms, the discrete approximation in terms of the percentage change between any two periods may be used as in Jorgenson and Griliches (1967).

Unlike other two approaches, the non-parametric approach requires extensive price information. In practice, the factor income shares obtained from secondary sources of data are often not reliable.

4.2 Measuring Total Factor Productivity using Indexes

4.2.1 Multifactor Productivity Indexes [2]

In a single-input single-output case, where outputs y_0 and y_1 are produced from inputs x_0 and x_1 respectively in period 0 and 1 respectively, the productivity at period 0 and period 1 will be,

$$\Pi_0 = y_0 / x_0 \tag{4.29}$$

$$\Pi_1 = y_1 / x_1 \tag{4.30}$$

[2] The illustration for multifactor productivity indexes shown in this section and afterwards is taken from Ray (2004)

so that the productivity index in period 1 with period 0 as the base is,

$$\pi_1 = \Pi_1 / \Pi_0 = (y_1 / x_1) / (y_0 / x_0) = (y_1 / y_0) / (x_1 / x_0) \qquad (4.31)$$

In a multiple-input multiple-output case, one must replace the simple ratios of the output and input quantities of equation 4.31 by quantity indexes of outputs and inputs, so that the multifactor productivity index is,

$$\pi_1 \quad = \Pi_1 / \Pi_0 \quad = Q_y / Q_x \qquad (4.32)$$

where Q_y and Q_x are respectively output and input quantity indexes of the firm in period *1* with period *0* respectively as the base.

Hence, different methods of indexing outputs and inputs will obtain different measures of multifactor productivity. The most common indexes for measuring productivity are described below.

4.2.2 Törnqvist Productivity Index

Törnqvist index is measured using weighted geometric means of the relative quantities from the two periods. The Törnqvist indexes for *m* output quantities and *n* input quantities in period 1 with respect to period 0 are,

$$TQ_y = \prod_{i=1}^{m} (y_1^i / y_0^i)^{vi} \qquad (4.33)$$

$$TQ_x = \prod_{j=1}^{n} (x_1^j / x_0^j)^{sj} \qquad (4.34)$$

where v^i is the arithmetic mean of v_0^i and v_1^i and s^j is the arithmetic mean of s_0^j and s_1^j and

$$v_t^{\,i} = (p_t^{\,i}\, y_t^{\,i}) \,/\, (\sum_{i=1}^{m} p_t^{\,i}\, y_t^{\,i}) \tag{4.35}$$

$$s_t^{\,j} = (w_t^{\,j}\, x_t^{\,j}) \,/\, (\sum_{j=1}^{n} w_t^{\,j}\, x_t^{\,j}) \tag{4.36}$$

where $p_t^{\,i}$ is the price of output y^i at time t and $w_t^{\,j}$ is the price of input x^j at time t.

The Törnqvist productivity index is the ratio of the Törnqvist output and input quantity indexes,

$$\pi_{TQ} = TQ_y \,/\, TQ_x \tag{4.37}$$

Törnqvist productivity index can be measured without any knowledge of the underlying technology, but it requires extensive price information along with quantity data.

4.2.3 Fisher Productivity Index

Laspeyres output quantity index is the ratio of the two output vectors at base prices,

$$LQ_y = (\sum_{i=1}^{m} p_0^{\,i}\, y_1^{\,i}) \,/\, (\sum_{i=1}^{m} p_0^{\,i}\, y_0^{\,i}) \tag{4.38}$$

Paasche output quantity index is the ratio of the two output vectors at current prices,

$$PQ_y = (\sum_{i=1}^{m} p_1^{\,i}\, y_1^{\,i}) \,/\, (\sum_{i=1}^{m} p_1^{\,i}\, y_0^{\,i}) \tag{4.39}$$

The Fisher output quantity index is the geometric mean of the Laspeyres and Paasche output quantity indexes,

$$FQ_y = (LQ_y \,.\, PQ_y)^{1/2} \tag{4.40}$$

Laspeyres input quantity index is the ratio of the two input vectors at base prices,

$$LQ_x = (\sum_{j=1}^{n} w_0^{\,i}\, x_1^{\,i}) \,/\, (\sum_{j=1}^{n} w_0^{\,i}\, x_0^{\,i}) \tag{4.41}$$

Paasche input quantity index is the ratio of the two input vectors at current prices,

$$PQ_x = (\sum_{j=1}^{n} w_1^{\,i}\, x_1^{\,i}) \,/\, (\sum_{j=1}^{n} w_1^{\,i}\, x_0^{\,i}) \tag{4.42}$$

The Fisher input quantity index is the geometric mean of the Laspeyres and Paasche input quantity indexes,

$$FQ_x = (LQ_x \,.\, PQ_x)^{1/2} \tag{4.43}$$

The Fisher productivity index is the ratio of the Fisher output and input quantity indexes,

$$\pi_{FQ} = FQ_y \,/\, FQ_x \tag{4.44}$$

Fisher productivity index can be measured without any knowledge of the underlying technology, but it requires extensive price information along with quantity data.

4.3 Measuring Productivity using Efficiency Measurement Method [3]

4.3.1 Production Possibility Set

Let a production technology utilize a vector of inputs, denoted $x = (x_1, \ldots\ldots\ldots, x_n)$ ε $\Re_+^{\,n}$ to produce a non-negative vector of outputs, denoted $y = (y_1, \ldots\ldots\ldots, y_m)$ ε $\Re_+^{\,m}$.

The production possibility set of a production unit (PU) is a subset T of the space $\Re_+^{\,m+n}$. A PU may select any input-output configuration (x, y) ε T as its production plan. The production possibility set is the collection of all feasible input and output vectors. It is represented as

$$T = \{(y, x) : \text{x can produce y}\} \text{ c } \Re_+^{\,m+n}$$

Furthermore, production possibility set can be represented by input requirement set $L(y)$ or output producible set $P(x)$.

The input requirement set is the collection of all input vectors $x = (x_1, \ldots\ldots\ldots, x_n)$ ε $\Re_+^{\,n}$ that yield at least output vector $y = (y_1, \ldots\ldots\ldots, y_n)$ ε $\Re_+^{\,m}$. It can be represented as

$$L(y) = \{x : (x, y) \text{ is feasible}\}$$

The output producible set is the collection of all output vectors $y = (y_1, \ldots\ldots\ldots, y_n)$ ε $\Re_+^{\,m}$ that are produced from the given input vector $x = (x_1, \ldots\ldots\ldots, x_n)$ ε $\Re_+^{\,n}$. It can be represented as

$$P(x) = \{y : (x, y) \text{ is feasible}\}$$

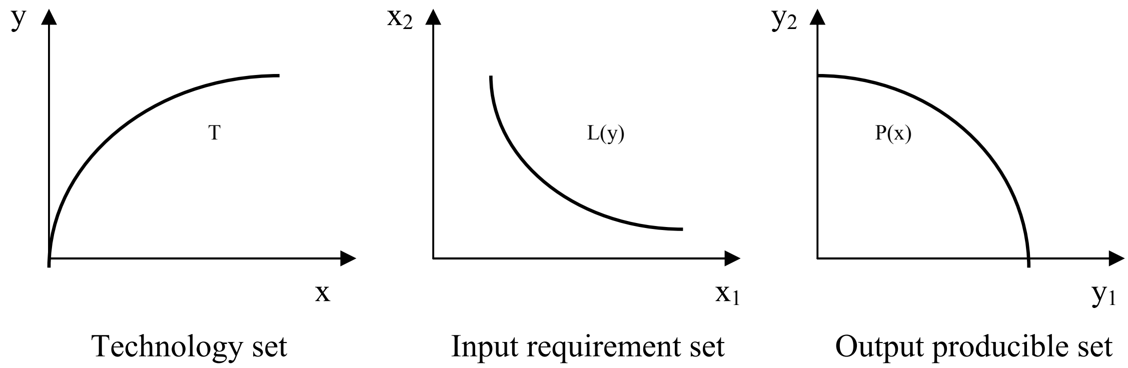
These production possibility sets are illustrated in Figure 4.1.

---

[3] The illustration for production possibility set and production frontiers shown in this section is taken from Fried et al. (1993)

Figure 4.1

Production Possibility Sets



4.3.2 Production Frontiers

To illustrate the concept of production frontier, one can use an important class of technologies having a single output *y* and an *n*-dimensional vector of input *x*. Suppose the production possibility set satisfies $T(x, y) \geq 0$. A general representation of the frontier technology is given as $y = f(x)$.

The function $f(.)$ is the production frontier and $y = f(x)$ gives the upper boundary of *T*. Given input *x*, the maximum producible output $y = f(x)$ can be achieved. In the form of maximization, the producible frontier is expressed by

$$f(x) = \max \{y' : T(x, y') \geq 0\}$$

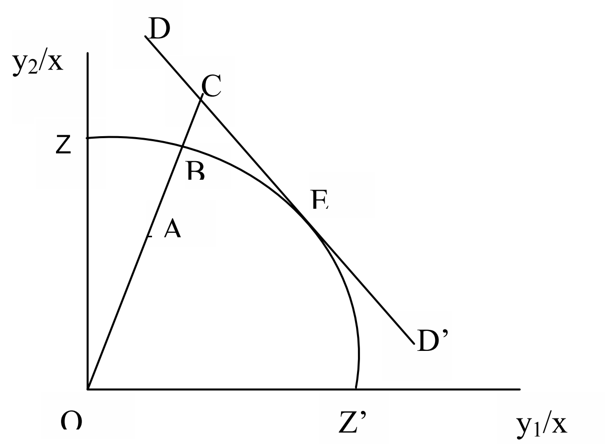
The production frontier serves as a standard against which to measure the technical efficiency of production. Only the efficient observations lie on this frontier, while the inefficient observations lie below the frontier. Production frontier has a property of scale economies: constant returns to scale, decreasing returns to scale and increasing returns to scale. The assumption of scale property is often required in production frontier estimation and efficiency analysis, especially in non-parametric frontier methods.

4.3.3 Efficiency

The usual measure of efficiency is the ratio of output to input. This measure is often inadequate due to the existence of multiple inputs and multiple outputs related to different resources, activities and environmental factors. Farrell (1957) first defined a simple measure of firm efficiency defining technical, allocative and total economic efficiency of the firm.

At this point, it should be noted that efficiency measurement may be input-orientated where the production unit seeks to minimize use of inputs for a fixed level of output or it may be output-orientated where the unit seeks to maximize production of outputs for a fixed level of input factors. The Farrell output-orientated efficiency measurement is presented in Figure 4.2.

Figure 4.2:
Output-Orientated Measurement of Technical, Allocative and Total Economic Efficiency



In Figure 4.2, the production unit at point *A* is producing two outputs (y_1 and y_2) from a single input (x). The line *ZZ'* is the production possibility curve. Here, the firm operating at *A* will be inefficient, as it is below the curve *ZZ'* representing the upper bound of production possibilities. Point *B*, being on the production possibility curve, is fully efficient and represents the efficient reference of point *A*. Hence, the

Technical Efficiency (TE) of a production unit operating at *A* will be measured as the ratio *OA/OB*.

$$TE(A) = OA/OB$$

This technical efficiency term denotes the ability of the firm to produce maximum output from a given amount of input factors.

If the price information is known, then the isorevenue line *DD'* can be drawn. This line will represent the combination of outputs each of which are leading to maximum amount of revenue. Allocative efficiency (AE) would be the ratio *OB/OC*, since the distance *CB* represents the increase in revenue that would occur if production unit was at the allocatively and technically efficient point *E*, instead of the technically efficient but allocatively inefficient point *B*.

$$AE(A) = OB/OC$$

This allocative efficiency reflects the ability of a firm to produce its outputs in optimal proportions, given their respective prices and the production technology. In other words, it is concerned with choosing between the different technically efficient combinations of outputs to be produced from the minimum possible usage of the input.

Total Economic Efficiency (EE) is the combination of technical and allocative efficiencies. This is the product of the technical efficiency and the allocative efficiency. Overall economic efficiency in this case would be *OA/OC*.

$$EE(A) = TE(A).AE(A) = (OA/OB).(OB/OC) = OA/OC$$

An organization will only be cost efficient if it is both technically and allocatively efficient.

So, the fundamental assumption of Farrell's (1957) concept was the possibility of inefficient operations, immediately pointing to a 'frontier production function' concept as the benchmark, as opposed to a notion of 'average performance' underlying most of the econometric literature.

4.3.4 Data Envelopment Analysis

The Malmquist productivity index is, in most of the cases, constructed using Data Envelopment Analysis (DEA) in the productivity literature. Hence, a brief description of DEA method is provided here before going on to describe the Malmquist productivity index calculations.

Like Farrell (1957), Charnes, Cooper and Rhodes (1978) too proposed efficiency measures and the framework of a piecewise linear production technology, but this linear programming (LP) model (CCR model of DEA) formulated was a generic one compared to Farrell's unit isoquant approach in the case of a single output. This CCR model was readily computable, either using standard LP codes or developing more efficient tailor-made software. One unique contribution of CCR is the explicit connection made between a productivity index in the form of a weighted sum of outputs on a weighted sum of inputs and the Farrell technical efficiency measure in the case of constant returns to scale. The starting point in CCR was to find weights by maximization of such a productivity ratio subject to best practice and normalization constraints.

DEA is a linear-programming methodology, which uses data on the input and output quantities of a group of administrative divisions to construct a piece-wise linear surface over the data points. This frontier surface is constructed by the solution of a sequence of linear programming problems – one for each division in the sample. The degree of technical inefficiency of each division, which is the distance between the observed data point and the frontier, is produced as a byproduct of the frontier construction method.

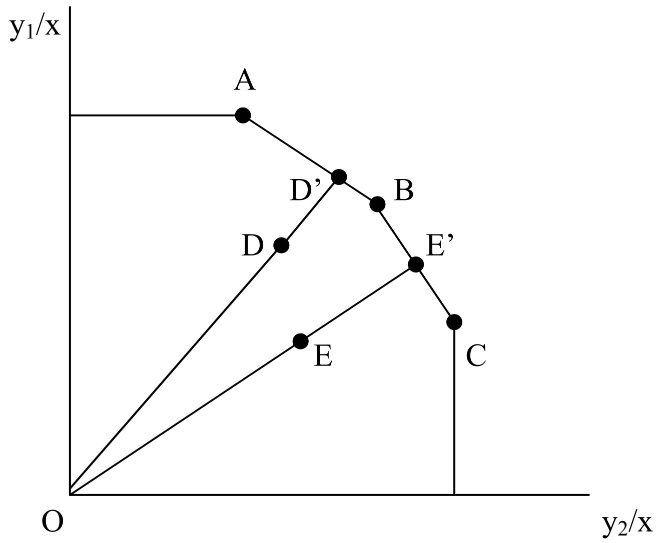
The DEA problem can be illustrated graphically with an example[4], where a group of five divisions is producing two outputs from a single input vector. These five divisions are depicted in Figure 4.3. Divisions *A*, *B* and *C* are efficient divisions because they define the frontier. Divisions *D* and *E* are inefficient divisions.

---

[4] The example is constructed following the illustration in Coelli and Rao (2003)

Figure 4.3

Illustration of an output orientated DEA model



For division *D* the technical efficiency score is $TE_D = OD/OD'$ and its peers are divisions *A* and *B*. In the DEA output listing this division would have non-zero *λ*-weights associated with divisions *A* and *B*. On the other hand, the DEA output listing for divisions *A*, *B* and *C* would provide technical efficiency scores equal to one and each division would be its own peer.

For *N* administrative regions or divisions in a particular time period, the linear programming problem solved for the *i*-th division in an output-orientated DEA model is as follows.

$$\max_{\varphi,\lambda} \quad \varphi$$
$$\text{subject to,} \quad -\varphi y_i + Y\lambda \geq 0,$$
$$xi - X\lambda \geq 0,$$
$$\lambda \geq 0, \quad (1)$$

where y_i is a $M \times 1$ vector of output quantities for the *i*-th division, x_i is a $K \times 1$ vector of input quantities for the *i*-th division, Y is a $M \times N$ matrix of output quantities for all N divisions, X is a $K \times N$ matrix of input quantities for all N divisions, λ is a $N \times 1$ vector of weights, and φ is a scalar.

It can be observed that, φ will take a value greater than or equal to one, and that φ-1 is the proportional increase in outputs that could be achieved by the *i*-th division, with input quantities held constant. Also, it can be noted that $1/\varphi$ defines a technical efficiency (TE) score which varies between zero and one.

The above linear program is solved *N* times, once for each division in the sample. Each LP produces one φ and one λ vector. The φ-parameter provides information on the technical efficiency score for the *i*-th division and the λ-vector provides information on the peers of the (inefficient) *i*-th division. The peers of the *i*-th division are those efficient divisions that define the facet of the frontier against which the (inefficient) *i*-th division is projected.

4.3.5 Malmquist Productivity Index

Caves, Christiensen and Deiwert (1982a,b) defined a Malmquist index which is the ratio of two output distance functions with the numerator being the output distance function at time *t*+1 using time *t* technology and the denominator being the output distance function at time *t* based on time *t* technology. An alternative is to define the distance functions in terms of technology in time *t*+1. Färe et al. (1994) defined a Malmquist index that is the geometric mean of these two indexes and this allowed them to interpret two terms in the index, one as efficiency changes represented by movements toward the frontier and the second as changes in technology represented by shifts in the frontier. This Malmquist index measures the productivity change between two data points.

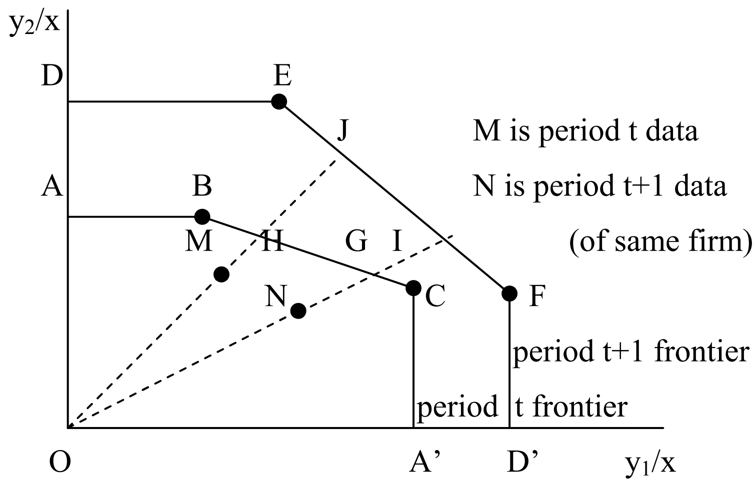
It can be illustrated graphically[5] how the Malmquist productivity index measures the productivity change considering a set of firms that use a single input *x* to produce a two outputs y_1 and y_2. It is assumed that, efficient operation of the firms is characterized by constant returns to scale and a perfect competition. The illustration of an output orientated Malmquist productivity index is provided in Figure 4.4.

---

[5] This output-orientated illustration follows the argument of the input-orientated illustration presented in Thanassoulis (2001)

Figure 4.4

Illustration of MPI in an Output-Orientated Case



Here, the same firm is operating at *M* in period *t* and at *N* in period *t*+1. The efficient frontier in period *t* is constructed as a piecewise linear production frontier *ABCA'* defined by fully efficient firms *A*, *B*, *C* and *A'* at time *t*. In period *t*+1, the piecewise linear frontier *DEFD'* is defined by fully efficient firms *D*, *E*, *F* and *D'* at time *t*+1.

If the frontier at period *t* had not moved at time *t*+1, the output efficiency of the firm at *M* at time *t* would be *OM/OH* and at time *t*+1 it would be *ON/OG*. So, here the ratio (*ON/OG*)/(*OM/OH*) equal to one would mean that the firm has stayed the same distance from boundary in periods *t* and *t*+1 and so there is no change in productivity.

If the same ratio is more than one, it would mean that there has been a progress in productivity, as the firm moved closer to the stationary efficient boundary in period *t*+1 than it was in period *t*. The same ratio after considering the period *t*+1 frontier as stationary would be (*ON/OI*)/(*OM/OJ*).

So in a situation where not only the firm moves relative to the efficient boundary over time but also the boundary itself moves, one has to compute the ratio once relative to the *t* period boundary and again relative to the *t*+1 period boundary. And, the productivity change of this firm in Figure 4.4 can be expressed as a geometric mean of the change with respect to period *t* frontier and the change with respect to period t+1 frontier. This expression is termed as the Malmquist Productivity Index (MPI), representing the change in productivity for the firm in question.

$$MPI = [\{(ON / OG) / (OM / OH)\} \times \{(ON / OI) / (OM / OJ)\}]^{0.5} \qquad (4.45)$$

Upon rearranging[6], the equation 4.45 can be expressed as equation 4.46.

$$MPI = [(ON / OI) / (OM / OH)] \times [(OI / OG) \,.\, (OJ / OH)]^{0.5} \qquad (4.46)$$

Here, *ON/OI* measures efficiency of the firm at time t+1 with respect to the boundary at time t+1, *OM/OH* measures efficiency of the firm at time t with respect to the boundary at time t. The first part of the index, i.e. the ratio of the efficiency at time t+1 to the efficiency at time t, is referred to as the catch-up term. The catch-up term measures how much closer to the boundary the firm is in period t+1 compared to period t. If the catch-up term is 1, the firm has the same distance in both periods from the respective frontiers. If the catch-up term is over 1, the firm has moved closer to the frontier in later period indicating increase in firm efficiency. If the catch-up term is below 1, the firm has moved away from the frontier in later period indicating decrease in firm efficiency.

[6] The steps for rearrangement of the terms from equation 4.45 to equation 4.6 are shown below.

Step-1, $MPI = [\{(ON / OG) / (OM / OH)\} \times \{(ON / OI) / (OM / OJ)\}]^{0.5}$

Step-2, $MPI = [(ON^2 \,.\, OH \,.\, OJ) / (OG \,.\, OM^2 \,.\, OI)]^{0.5}$

Step-3, $MPI = [ON / OM] \,.\, [(OH \,.\, OJ) / (OG \,.\, OI)]^{0.5}$

Step-4, ,$MPI = [ON / OM] \,.\, [(OH^2 \,.\, OJ \,.\, OI) / (OH \,.\, OG \,.\, OI^2)]^{0.5}$

Step-5, $MPI = [(ON \,.\, OH) / (OM \,.\, OI)] \,.\, [(OJ \,.\, OI) / (OH \,.\, OG)]^{0.5}$

Step-6, $MPI = [(ON / OI) / (OM / OH)] \times [(OI / OG) \,.\, (OJ / OH)]^{0.5}$

In equation 4.46, the term *OI/OG* measures the distance of two boundaries at the output mix of the firm in period t+1, whereas *OJ/OH* measures the distance of two boundaries at the output mix of the firm in period t. The last part of the expression, the geometric mean of these two distances, is referred to as the boundary-shift term. The boundary-shift term measures the movement of the frontier itself between period t and t+1 at two locations. If the boundary-shift term is more than 1, this represents productivity gain by the industry in that the firm has higher output levels in period t+1 than in period t at both output mixes considered for the standard unit of input. If the boundary-shift term is equal to 1, either the t and t+1 boundaries would have been coincident at the two output mixes involved, or at one output mix outputs in period t would have exceeded those in period t+1 and the converse would have been true at the other so that on an average the industry boundary would have shown neither gain nor loss in productivity. Finally, a boundary shift term under *1* would signal that the industry has registered productivity loss as output levels would on average be lower in period t+1 compared to period t, controlling for the input level.

Following Färe et al (1994), the output-orientated Malmquist index for every division between period s (the base period) and period t (the recent period) is given by

$$M(y_s, x_s, y_t, x_t) = \left[\frac{d^s(y_t, x_t)}{d^s(y_s, x_s)} \times \frac{d^t(y_t, x_t)}{d^t(y_s, x_s)} \right]^{1/2} \tag{4.47}$$

where $d^s(y_t, x_t)$ is the distance from the period t observation to the period s technology, $d^s(y_s, x_s)$ is the distance from the period s observation to the period s technology, $d^t(y_s, x_s)$ is the distance from the period s observation to the period t technology, and $d^t(y_t, x_t)$ is the distance from the period t observation to the period t technology.

Equation 4.47 can be expressed allowing for a decomposition as equation 4.48.

$$M(y_s, x_s, y_t, x_t) = \frac{d^t(y_t, x_t)}{d^s(y_s, x_s)} \left[\frac{d^s(y_t, x_t)}{d^t(y_t, x_t)} \times \frac{d^s(y_s, x_s)}{d^t(y_s, x_s)} \right]^{1/2} \tag{4.48}$$

Here, the first term of the product at the right hand side measures the change in the output-orientated measure of Farrell technical efficiency between periods *s* and *t*. That is, the efficiency change is equivalent to the ratio of the technical efficiency in period *t* to the technical efficiency in period *s*. The remaining part of the index in equation 4.48 is a measure of technical change. It is the geometric mean of the shift in technology between the two periods, evaluated at x_t and also at x_s.

Both Törnqvist productivity index and Fisher productivity index can be measured without any knowledge of the underlying technology, but both requires price information along with quantity data. By contrast, the Malmquist productivity index is a normative measure that constructs a production frontier representing the technology and uses the corresponding distance functions evaluated at different input-output combinations for productivity comparison. There are several advantages of this approach. First, it allows for inefficiency and enables decomposition of the index into production unit's technical efficiency change component and industry wide technical change component. Second, it does not require any price data for productivity assessment. Third, it does not require economic estimation and can be implemented with a data envelopment technique.

If inefficiency is assumed away, the Malmquist index features only the technical change component which corresponds to technical change as defined by Solow under the assumption of competitive behavior and as measured by the conventional indices such as Törnqvist and Fisher indices (Raa and Sestalova, 2006).

4.3.6 Interpretation of Malmquist Productivity Index in terms of Total Factor Productivity Growth [7]

Let us consider the case of one output and neutral technical changes. In this case, the technology can be represented by a production function of the following form.

$$y_t = A(t)\, f(x_t) \tag{4.49}$$

[7] This illustration for interpretation of Malmquist Productivity Index in terms of Total Factor Productivity Growth is taken from Raa and Sestalova (2006)

where, y_t is output at period t, x_t is input at period t, $A(t)$ is the total factor productivity at period t. Hence, the Solow Residual between period t and period s in the discrete case can be expressed as the technical change component,

$$TFPG = lnA(s) - lnA(t) = ln\ [A(s) / A(t)] \qquad (4.50)$$

It can be observed that in this special case that the condition of optimizing behavior i.e. assumption of no inefficiency yields the equivalence of the technical change component (boundary-shift term) of Malmquist Productivity Index. In particular, for this production function the output distance function at any particular period T will be as follows:

$$D^T (y_T, x_T) = y_T / \{A(T) \cdot f(x_T)\} \qquad (4.51)$$

Substituting this into the formula for the Malmquist index in equation 4.47 yields:

$$M (y_s, x_s, y_t, x_t) = [y_s / f(x_s)] / [y_t / f(x_t)] \qquad (4.52)$$

Assuming inefficiency away, we obtain that output and input in each time t are related by equation 4.49, which yields the result:

$$M (y_s, x_s, y_t, x_t) = A(s) / A(t) \qquad (4.53)$$

Hence the technical change component of Malmquist index is equivalent to the Solow residual measure of technical change equation 4.50 above.