Business Performance Measurement and Management

Business Performance Measurement and Management

Edited by

Vincent Charles and Mukesh Kumar

CAMBRIDGE SCHOLARS

PUBLISHING

Business Performance Measurement and Management,
Edited by Vincent Charles and Mukesh Kumar

This book first published 2014

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-6139-1, ISBN (13): 978-1-4438-6139-7

CONTENTS

PREFACE

Measuring and managing the performance of a business is one of the main requirements of the management of any organization. *Performance management* is a broad approach to planning, measuring, and monitoring the company's business activities. It focuses not only on individual employees but also on teams, programmes, processes, and the organisation as a whole. An effective business performance management framework enables businesses to define strategic goals and then measure and manage performance against these goals. Neely, Adams, and Kennerley (2002) defined *performance measurement* as the process of quantifying the efficiency and effectiveness of past action. However, it does not reveal the process that individual managers went through in setting the initial targets, the actions that were going to be required, the anticipated state of the business environment for which those actions were conceived, whether or not the required actions were actually carried out, and whether those actions actually contributed to the success. Without this knowledge, measures are at best misleading and, in the worst case, will promote responses that are ill-considered and damaging to the long-term prospects of the organization.

This book introduces new contexts and themes of application and presents emerging research areas related to business performance measurement and management. It draws authors from a variety of functional disciplines, all of whom are working in the field of business performance measurement and management, and thus resulting in a variety of perspectives on performance measurement from various functional areas – accounting, finance, economics, marketing, and operations management – in a single volume. The book, titled *Business Performance Measurement and Management,* is well organized into 22 chapters contributed by researchers from all around the globe and covering a range of issues consisting of conceptual issues, applications, and theoretical contributions related to performance management in business.

Chapter One surveys the different methods of total factor productivity, an economic performance measurement tool. A brief overview of non-parametric and parametric methods under both a non-frontier approach (which ignores efficiency) and a frontier approach (which explicitly allows for inefficiency) has been provided, enumerating their relative merits and demerits.

Chapter Two explores the empirical studies on green supply chain activities and develops a performance measurement framework consisting of environmental, economic, and social performance metrics, which serves as a practical platform for decision makers.

Chapter Three brings together some of the main scholarly sources of corporate issues linked to corporate citizenship in the CSR discussion, which is particularly important in today's global business.

Chapter Four examines the current changes in the business environment for management education and how these changes are influencing transformation of management education. It also highlights some possible ways which may assist in addressing these challenges.

Chapter Five aims to develop correlational parameters and maturity indicators in the context of higher education in India by means of an extensive opinion survey of stakeholders of institutions of higher education and parameterized rating and uses these indicators to filter the number of institutions for further intense study. This study could be helpful to the institutions of higher education that are struggling to cope with the variable market dynamics and are planning to transform their organizations.

Chapter Six presents the results of the application of the enterprise expert search system to the tasks introduced at the Text Retrieval Conference (TREC). Two specific indicators are used in order to treat the lexicon statistically: (a) calculating lexicon-candidate connection power reveals definite terms which are characteristic for a candidate so this candidate can be found by such terms and (b) calculating the weight of the lexicon allows extraction from the whole collection of a small portion of vocabulary, named as significant. The significant lexicon enables an effective search to be performed in thematically specialized knowledge fields. So the search engine minimizes the lexicon necessary for answering a query by extracting the most important part from it.

Chapter Seven presents the abstract of the book *Learning with Lean: Unleashing the Potential for Sustainable Competitive Advantage*, Taylor & Francis, New York NY, 2013. ISBN-13: 978-1466572949.

Chapter Eight assesses the repercussions of the financial crisis on the training budgets and practices of key government entities in the Emirate of Dubai in the United Arab Emirates (UAE). It examines the alternative approaches the government introduced and implemented to cope with diminishing training budgets, and it assesses their effectiveness. The chapter concludes by providing strategic recommendations aimed at guiding the government of Dubai and other governments in the region to improve the quality of their training programs during times of financial constraints.

Chapter Nine explains the benefits of the alternative decomposition of return on equity (ROE) with the help of a case. A company is selected from Istanbul Stock Exchange (ISE 100) and the company's ROE is calculated according to two approaches of decomposition of ROE. At the end of the case, it is concluded that the company is unable to manage the financing activities successfully and thus the financing activities result in a decrease in ROE. In order to increase the ROE, the company should borrow at a lower rate or decrease the level of financial leverage. Since this information is not provided through the standard DuPont Analysis, it is concluded that the alternative decomposition of ROE is more useful to develop corporate strategies.

Chapter Ten is concerned with the quality and delivery of a primary healthcare facility in a developing economy like India where "advancements of a few pockets are highlighted while the sub-human conditions of others just do not find any avenue for a decent living" (Sengupta & Mukherjee, 2010, p. 558). Traditional analyses point to market failure, which may be corrected by government intervention. However, the government may fail to deliver, leading to consideration of the concept of public private partnership (PPP). The authors have tried to conceptualize this within a rigorous framework, demonstrating wide inequality, market exclusion, government failure, and justification of PPP.

Chapter Eleven provides a practical application on the utilization of a benchmarking process adopted by South Africa's electric utility, Eskom, in its pursuance of the four tenets used in any production business: the accessibility to the product, the availability of the product, its reliability, and its "better value for money" or affordability.

Chapter Twelve presents the results of a study on management functions in dual-purpose cattle farming systems located in the municipalities of Catatumbo and Colón in Venezuela. These functions are studied by defining and calculating synthetic management indices that collect information on the behaviour of the managers of farms in the area under study.

Chapter Thirteen presents a model for the acceptance of business processes by employees. In this context, the authors developed an authentic questionnaire to collect data from people who are interacting with certain process-focused models and standards used for improvements of systems and software engineering and management business processes. The application of partial least square structural equation modelling resulted in developing the model with 18 imperative factors and their statistically significant relationship. Furthermore, the authors developed a checklist to test and promote the acceptance of business processes. Both the model and the pertinent checklist are truly beneficial for business process definition, deployment, implementation, and maintenance activities related to systems and software engineering and management.

Chapter Fourteen proposes four conceptual frameworks employing the same six constructs (namely, service quality, trust, switching cost, corporate image, customer satisfaction, and customer loyalty) to examine which model explains mobile subscribers' loyalty in the best possible way for a leading mobile operator in Bangladesh.

Chapter Fifteen aims to examine the extent to which AC Milan could improve its payoff, following the optimal strategies derived based on match statistics collected from the UEFA Champions League game between AC Milan and FC Barcelona through the application of some deterministic, possibility, stochastic, and fuzzy LP models.

By means of a stochastic approximation, Chapter Sixteen proposes to estimate the necessary design parameters within a range of desired accuracy for a given target value for the performance function. The proposed solution algorithm is based on Newton's methods, using a single-run simulation to minimize a loss function that measures the deviation from a target value. The properties of the solution algorithm and the validity of the estimates are examined by applying them to reliability and queuing systems with a known analytical solution.

Chapter Seventeen uses the mixed non-linear integer programming (MNLIP) model to examine whether wage differences between Super talent and Normal players improve the performance of four teams, which participate in a tournament such as in the UEFA Champions League (UCL) group matches. With ad hoc wage differences, the optimal solutions of the model show that higher wage equality seems to improve the performance of all teams, irrespective of whether the elasticity of substitution between Super- and Normal- players is high or low.

Chapter Eighteen provides empirical evidence on the relationship between firm environmental performance and economic performance in two U.S. industries that are typically viewed as "highly environmentally sensitive" and the S&P 500 firms. The results reveal that firms that rated as high on environmental strengths have a higher economic performance than firms that ranked as low. This implies that investing resources to improve an organization's environmental performance can have a positive impact on its economic performance.

Chapter Nineteen aims to develop a method permitting simultaneous measure of technical and allocative efficiencies by introducing some argumentative modifications into the model structure that was developed by Charnes, Cooper, and Rhodes in 1978 (CCR).

Chapter Twenty aims to assess the relative efficiency of the 50 U.S. states, as well as estimate for each of them feasible reductions in taxes, debt, and public expenditures, by applying a two-stage network data envelopment analysis (DEA) approach for the period 2007-2011. The results reveal that, on average, states governed by the Democratic Party showed greater inefficiencies relative to GDP than those governed by the Republican Party.

Chapter Twenty-one proposes an integrated approach to the data envelopment analysis (DEA) and analytic hierarchy process (AHP) methodologies to overcome the problematic issue of confronting the contradiction between the efficiency and effectiveness of decision-making units (DMUs). A parametric goal programming model, with normalized data, has been developed in order to minimize the deviations between weights in DEA and target weights as computed by an AHP. By varying a parameter within a domain of efficiency losses, the author explores the potential trade-off that may exist between efficiency and effectiveness. This may result in different ranking positions of DMUs. An illustrative example, with synthesized data, is used to highlight the usefulness of the proposed approach.

Chapter Twenty-two presents an integrated model based on the analytic hierarchy process (AHP) and data envelopment analysis (DEA) methodologies that can be used to extract benefits from both methods by reflecting the priority weights of financial ratios in assessing the efficiency value of stocks. In the first stage, the priority weights of financial and market ratios are computed by AHP. In the second stage, by using a weighted average approach, the priority weights are integrated in the Andersen-Petersen (AP) model under conditions of variable returns to scale (VRS). An illustrated example of eight listed companies in the steel industry of China is used to highlight the usefulness of the proposed model.

The chapters contributed to this book should be of considerable interest and provide readers with informative reading.

ACKNOWLEDGEMENTS

The many academics and researchers who contributed articles and the experts within the field of business performance measurement and management who reviewed the articles made this book possible. We thank you.

LIST OF REVIEWERS

1. Alan Marchant, Utah State University, Logan, USA
2. Andrej Grebenšek, University of Ljubljana, Portorož, Slovenia
3. Alexander Gromoff, National Research University Higher School Of Economics, Moscow, Russia
4. Angel Prieto, Instituto de Recursos Naturales y Agrobiología, Consejo Superior de Investigaciones Científicas, Salamanca, Spain.
5. Asli Türel, Istanbul University, Istanbul, Turkey
6. Ayub Mehar, Iqra University, Karachi, Pakistan
7. Belaïd Aouni, Laurentian University, Sudbury, Canada
8. Bénédicte Le Grand, Université Paris 1 Panthéon-Sorbonne, Paris, France
9. Biresh K. Sahoo, Xavier Institute of Management, Bhubaneswar, India
10. Boris Slavin, Center for IT expertise of The Union of CIOs of Russia, Moscow, Russia
11. David Vivas-Consuelo, Universidad Politécnica de Valencia, Valencia, Spain
12. Ezutah Udoncy Olugu, University of Malaya, Kuala Lumpur, Malaysia
13. Frederick T. Evers, University of Guelph, Guelph, Canada
14. Gabriel Villa Caro, Universidad de Sevilla, Sevilla, Spain
15. Hilary Bradbury Huang, Oregon Health & Science University, Portland, USA
16. Hossein Azizi, Islamic Azad University, Parsabad Moghan, Iran
17. John S. Liu, Graduate Institute of Technology Management, Taipei, Taiwan
18. Juan Bautista Delgado García, Universidad de Burgos, Burgos, Spain
19. K. S. Sujit, Institute of Management Technology, Dubai, UAE
20. Katarína Čulková, Institute of Business and Management, Technical University of Košice, Košice, Slovakia
21. Laurie N. DiPadova-Stocks, Park University, Parkville, USA
22. María Caballer-Tarazona, University of Valencia, Valencia, Spain
23. Max Kleiman-Weiner, Massachusetts Institute of Technology, Cambridge, USA
24. Michael Lightfoot, London Knowledge Lab, Institute of Education, London, UK
25. Neha Purushottam, University of South Africa, Gauteng, South Africa

26. Plácido Rodríguez, Universidad de Oviedo, Gijón, Spain
27. Robert Breunig, Australian National University, Canberra, Australia
28. Sue R. Faerman, University at Albany, State University of New York, Albany, USA
29. Sungsoo Kim, Rutgers, The State University of New Jersey, Camden, USA
30. Swadhin Kumar Mondal, University of Delhi, Delhi, India
31. Thomas V. Huynh, Naval Postgraduate School, Monterey, USA
32. Xin Zhang, Austin Peay State University, Clarksville, USA

CHAPTER ONE

PERFORMANCE MEASUREMENT IN TERMS OF TOTAL FACTOR PRODUCTIVITY GROWTH: A SURVEY OF THE EVOLUTION OF DIFFERENT APPROACHES

MUKESH KUMAR AND VINCENT CHARLES

Abstract

In this chapter, the different methods of total factor productivity measurement are surveyed. A brief overview of non-parametric and parametric methods under both a non-frontier approach (which ignores efficiency) and a frontier approach (which explicitly allows for inefficiency) has been provided, with their relative merits and demerits.

1.1 Introduction

Productivity measures are frequently operationalised in terms of ratios of individual output to individual input, which is referred to as *partial factor productivity*. However, such productivity can be misleading in drawing any conclusion about the performance of the input. For example, an increase in the output per unit of labour may not necessarily be attributed to the increase in labour productivity because other inputs (capital, skilled workers, etc.) are used simultaneously in the production process.

Total factor productivity (TFP) growth is defined as the ratio of output to a weighted combination of inputs. Thus, it is a generalisation of partial factor productivity measures. TFP growth is of crucial significance in the context of economic growth, particularly in developing countries, as these economies are often faced with an acute shortage of productivity resources. The rate of industrial growth is determined by the rate of expansion of productive resources and the rate of growth in TFP, that is, the overall efficiency in the use of resources.

The different approaches to productivity measurement can be divided broadly into two groups: the frontier (modern) approach and the non-frontier (conventional) approach. Each one can further be subdivided into *parametric* and *non-parametric* methods. The non-frontier approach to productivity measurement is based on the assumption that the observed production in each period is equivalent to the production frontier, that is, the boundary of the technology is assumed to pass through the observed points, whereas, the frontier approaches explicitly account for inefficiency (Kumar & Basu, 2008).

This paper contains different subsections, which provide a brief overview of the different approaches to productivity measurement. Sections 1.2 and 1.3 summarise the non-frontier approaches, which ignore inefficiency and measure productivity growth either by means of non-parametric models or by means of index number methods and parametric models which use stochastic econometric methods. The next two subsections deal with the frontier approaches which explicitly allow for inefficiency. Sections 1.4 and 1.5, respectively, cover non-parametric and parametric frontier approaches to productivity measurement. Finally, in Section 1.6, the relative merits and demerits of different approaches to productivity measurement are highlighted.

1.2 Non-Parametric, Non-Frontier Approach

This method includes different index number approach and growth accounting models. The origin of the growth accounting approach to TFP growth can be traced to Tinbergen (1942) and Solow (1957). A number of alternative growth accounting estimates of TFP growth indices can be derived on the basis of alternative assumptions with respect to the underlying production function and common assumptions of competitive equilibrium and constant returns to scale (CRS).

The concept of TFP, defined as the ratio of real output to real input (a weighted sum of different inputs), was introduced by Tinbergen in 1942, while making an attempt to compare productivity growth among different countries.

Early generation of TFP studies generally used sets of *representative* input prices and output prices as weights for their respective inputs and outputs (Kendrick, 1961, 1973; Kendrick & Grossman, 1980). These measurements were variants of Laspeyres (1871) quantity indexes. For measurements over time, certain *base* periods have usually been chosen as a reference, while for cross-sectional measurements, certain production units usually have been selected as reference units. Criteria for the choice of such reference periods or units were mostly based on qualitative judgements.

A Laspeyres productivity index, TFP_L^1, measures TFP (at t = 1) as a ratio of Laspeyres output quantity index at t = 1 to a Laspeyres input quantity index at t = 1 with t = 0 as the reference (base):

$$TFP_L^1 = \frac{Q_L\left(P^0, P^1, Y^0, Y^1\right)}{Q_L\left(W^0, W^1, X^0, X^1\right)} = \frac{\left(P^0.Y^1\right)/\left(P^0.Y^0\right)}{\left(W^0.X^1\right)/\left(W^0.X^0\right)},$$

where Y and X are output and input and P and W indicate their prices, respectively.

Alternatively, Paasche TFP indexes, TFP_P^1, are analogous in their formulation to Laspeyres indexes except for their use of end period input and output prices, W^1 and P^1, as weights, that is,

$$TFP_P^1 = \frac{\left(P^1.Y^1\right)/\left(P^1.Y^0\right)}{\left(W^1.X^1\right)/\left(W^1.X^0\right)}.$$

Later on, Stigler (1941) developed the concept independently and suggested that a measure of real total factor input could be obtained by weighting inputs by their marginal products.

Solow (1957) provided an elementary way of segregating variations in output per head due to technical changes from those due to changes in the availability of capital per head. He defined *technological change* as a shorthand expression for any kind of shift in the production function. Assuming continuous time, Hicks' neutral technological change, the production function is taken as

$$y^t = A(t)\, f(x^t) \tag{1.3}$$

where $A(t)$ measures the cumulative effect of shifts over time. Again following Solow, let us assume that f is homogenous of degree 1, and inputs are paid the value of their marginal products, that is, $\partial f / \partial x_n^t = w_n^t / p^t$, where $w^t \in R_{++}^N$ is the vector of input prices in period t and $p^t \in R_{++}$ is the output price in period t. This assumption presumes that producers maximise the profit, implying no technical or allocative efficiency. The time deviations of the production function (1.3) give the growth accounting definition of productivity as

$$\frac{\dot{A}}{A} = \frac{\dot{y}}{y} - \sum_{n=1}^{N} S_n \left(\frac{\dot{x}_n}{x_n} \right) \tag{1.4}$$

where dots indicate time derivatives,

$$S_n = \frac{\partial y}{\partial x_n} \frac{x_n}{y} = \frac{w_n x_n}{py} = w_n x_n \Big/ \sum_{n=1}^{N} w_n x_n .$$

The continuous time formulation (1.4) is the residual growth in output not accounted for by growth in inputs associated with Solow (1957), Denison (1972), Kendrick (1961), and Jorgenson and Griliches (1967). In order to calculate the productivity (1.4), Solow made the assumption that the time derivatives could be approximated by discrete changes.

Kendrick's (1961) arithmetic measure approaches the measurement of productivity growth by means of using a distribution equation. He implicitly assumes a homogeneous production function and Euler's condition to obtain the following measure:

$$\frac{\dot{A}}{A} = \frac{y_1 / y_0}{(wL_1 + rK_1)/(wL_0 + rK_0)} \tag{1.5}$$

where w and r are the wage rate and rate of return on capital, respectively, variables with subscript 1 refer to the current period, and those with subscript 0 refer to the base period. The weights in this measure change over time and the aggregate production function consistent with this index is

$$y = \frac{tKL}{\left(cL^P + dK^P\right)^{1/P}} \tag{1.6}$$

which is a linear homogeneous production function with constant elasticity of substitution $\sigma = 1/(1+P)$; *c* and *d* are the efficiency parameters, *P* is the elasticity parameter, and *t* is the disembodied neutral technological change.

Under the assumption of competitive equilibrium, Kendrick's measure is equivalent to Solow's measure for small changes in the quantities of inputs and outputs.

The starting point for the derivation of Divisia (1926) indexes to TFP measure is the equality between total revenues *P.Y* and total cost *W.X*. It is assumed that input prices W and output prices P are unaffected by producers' input X and output Y decisions (i.e., the markets for all inputs and outputs are perfectly competitive). In competitive capital markets, the opportunity cost of capital equals the normal return on capital, and the above normal profits are zero. Differentiation of the equation *P.Y* = *W.X*, with respect to time, yields

$$\frac{\dot{TFP}_D}{TFP_D} = \sum_{j=1}^{m} \frac{\dot{y}_j}{y_j}\beta_j - \sum_{j=1}^{n} \frac{\dot{x}_i}{x_i}\alpha_i = \sum_{j=1}^{n} \frac{\dot{w}_i}{w_i}\alpha_i - \sum_{j=1}^{m} \frac{\dot{p}_j}{p_j}\beta_j \tag{1.7}$$

where $\beta_j = p_j y_j \Big/ \sum_{j=1}^{m} p_j y_j$ is the share of revenue generated by the *jth* output from total revenues and α_j is the share of the cost incurred by the *ith* input in total inputs. The percentage change in the TFP Divisia index is, thus, the difference between the sums of the weighted changes in outputs and inputs. This difference equals the difference between the sum of the weighted changes in quantities and prices.

If the continuous growth rates of Solow (1957), as defined in (1.4), are replaced by the discrete difference in logarithms, that is, $\frac{dy/y}{y} = \ln y^{t+1} - \ln y^t$ and input shares are calculated as an arithmetic mean, the index in (1.4) becomes equivalent to the Törnqvist (1936) index (TI) of TFP growth.

$$\ln TI = \ln y^{t+1} - \ln y^t - \sum_{n=1}^{N} \frac{1}{2}[S_n(t+1) + S_n(t)](\ln x_n^{t+1} - \ln x_n^t) \tag{1.8}$$

The Törnqvist index is *exact* if the technology in (1.3) is of translog form (Diewert, 1976). Since the (linearly homogenous) translog production function is flexible, that is, it is a second-order approximation to any arbitrary twice differentiable (linearly homogeneous) production function, the Törnqvist index is also a "superlative" index (Diewert, 1976).

Diewert (1992) examined the applicability of Fisher's price and quantity indexes to productivity measurement (Fisher, 1921). A Fisher TFP index, TFP_F, is the geometric mean of Laspeyres and Paasche TFP indexes, that is, $TFP_F = (TFP_L.TFP_P)^{1/2}$. TFP_F is shown by Diewert (1992) to be superlative by virtue of being exact for a flexible variable profit function.

1.3 Parametric, Non-Frontier Approaches

This section includes the estimation of the TFP growth by using the aggregate production function. The growth accounting models and the index number approaches, as discussed in the previous section, have the advantage of computational simplicity (there are no parameters to be estimated), but that is achieved at the cost of ignoring the measurement or sampling error. Thus, the resulting measures of productivity growth may be biased, and there is no notion of the precision with which productivity growth is measured.

An alternative approach is to parameterise the production function and estimate the parameters.

$$y^t = f\left(x^t, t\right) + \varepsilon$$

for $t = 1,2,...,T$. The estimated parameters are then used to solve for technological change as $\partial \ln f\left(x^t, t\right) / \partial t$. Given no change in the technical efficiency, this is equivalent to the TFP growth.

Some of the earlier studies used the production function approach to estimate the rate of technological progress. Gujarati (1967) used the Cobb-Douglas production function to assess the relative importance to capital, labour, and technology in explaining output growth in Indian manufacturing during 1946-1958. He found a significant favourable shift in the production function in only 8 out of the 28 industries studied. His estimates suggest that for 28 industries, taken together, the contribution of technological progress during the above period was rather small. Mehta (1976) carried out a similar exercise for the period 1953-1963 and arrived at a similar conclusion.

1.4 Non-Parametric Frontier Approach

Economics and operations research have common interests as to several research fields, one of the most prominent being the analysis of the production possibilities for micro units. The specific research stand of efficiency measurement for production units in the field of Operations Research was initiated with *Measuring the Efficiency of Decision Making Units* by Charnes, Cooper, and Rhodes (CCR) as the seminal paper in 1978.

Farrell (1957) laid the foundation for new approaches to efficiency and productivity studies at the micro level, involving new insights on two issues: how to define efficiency and productivity and how to calculate the benchmark technology and the efficiency measures.

The fundamental assumption was the possibility of inefficient operations, immediately pointing to a frontier production function concept as the benchmark, as opposed to the notion of average performance underlying the traditional approaches to production function estimations. Farrell's contribution was path-breaking in three aspects:

1. Efficiency measures were based on radial uniform contractions or expansions from inefficient observations to the frontier.
2. The production frontier was specified as the most pessimistic piecewise linear envelopment of the data.
3. The frontier was calculated through solving systems of linear equations, obeying the following two conditions: (i) that its slope is not positive and (ii) that no observed point lies between it and the origin.

It was Farrell (1957) who provided definitions and computational methods for both technical and allocative inefficiency, with the help of an unobserved production function (frontier), $y = f(x_1, x_2)$, which is characterised by a unit isoquant assuming CRS.

If the firm observed is using (x_1^0, x_2^0) to produce y^0, let point A in Figure 1.2 represent $(x_1^0/y^0, x_2^0/y^0)$. Then the ratio OB/OA gives the measure of technical inefficiency. Let PP' represent the isocost line which is the locus of combination of inputs to produce the unit output at the given input prices. The ratio *OD*/*OB* measures allocative inefficiency since the cost of point *D* is the same as that of the allocatively efficient point *C* and is less than that of the technically efficient point *B*. Lastly, the ratio *OD*/*OA* measures total inefficiency as a multiplication of technical and allocative inefficiency.

Charnes, Cooper, and Rhodes (1978) pioneered the technique of data envelopment analysis (DEA), a linear programming-based technique for measuring the relative performance/efficiency of the organisational units, where the presence of multiple inputs and outputs makes the comparison difficult. The CCR model primarily deals with non-linear (non-convex) programming, which is converted into equivalent linear programming to define a scalar increase of efficiency from the observed data on inputs and outputs. The efficiency measure defined in this fashion is equivalent to the productive efficiency defined by Farrell (1957). The relative efficiency score of a unit represents the maximum proportion of its inputs that the unit should have been using, if efficient, in order to secure at least its current output levels. Alternatively, the inverse of the efficiency score is the minimum factor, by which its inputs remain at their current levels.

The essential characteristic of the CCR formulation is the reduction of the multiple output-multiple input DMU situation to that of single virtual outputs and virtual inputs, for which the ratio of single virtual outputs to virtual inputs could be used to define the relative efficiency in a manner similar to that in engineering practice. Charnes, Cooper, Seiford, and Stutz (1982) developed a multiplicative DEA model by means of employing virtual outputs and inputs as in the CCR method to measure the relative efficiency where the resultant production function is piecewise log-linear rather than piecewise linear.

Banker, Charnes, Cooper, & Schinnar (1981) proposed the bi-extremal principle to locate efficiency frontiers and evaluate the efficiency of the DMUs, which can be accomplished from observational data by means of DEA, originally pioneered by Charnes et al. (1978). The bi-extremal principle, though non-linear, could be reducible to a finite sequence of linear programming problems. It has been illustrated by means of multiple output functions, which are piecewise Cobb-Douglas or general log linear type, and which allow for increasing, decreasing, and CRS.

Malmquist (1953) proposed the TFP indexes based on distance functions without the requirement of reliable data estimates of output and input prices. To conceptualise an output distance function, the technological frontier $F^t(X,Y)=0$ at time *t* can be represented by the input requirement function

$$x_1 = g^t(x_2,\ldots,x_n,Y),$$

where x_1 is the minimum amount of input 1 required to produce the vector of outputs Y, given the availability of the input quantities $x_2,\ldots,x_n$ for inputs $2,\ldots,n$. The output distance function g^t for $t=0$ is defined as

$$d^t(Y,X)=\max_{\delta}>0,$$

where $d^t(Y,X)$ is the maximum deflation factor δ^*, which will put the deflated output vector Y/δ^* and the input vector X on the production frontier. The distance δ^* can thus be interpreted as a measure of the maximal possible increase in technical efficiency, assuming that, in moving from a technically inefficient production to the production frontier, outputs would be scaled upward equiproportionately. Similarly, an input distance function is defined in terms of the maximal deflation factor that will just put the equiproportional deflated input vector X/δ and the output vector Y on the production frontier. Thus, the input distance function can be viewed as a measure of the maximal possible increase in technical efficiency, assuming that inputs would be scaled down equiproportionately in moving from a technically inefficient production to the production frontier.

The later concept can be traced to Farrell (1957), who measured technical efficiency by the maximal feasible proportional contraction in inputs.

Caves, Christensen, and Diewert (1982) defined two Malmquist output quantity indexes, $Q_M^O(Y^0,Y^1)$and $Q_M^1(Y^0,Y^1)$, as follows:

$$Q_M^O(Y^0,Y^1)=\frac{d^0(Y^1,X^0)}{d^0(Y^0,X^0)} \text{ and } Q_M^1(Y^0,Y^1)=\frac{d^0(Y^1,X^0)}{d^0(Y^0,X^0)}$$

These Malmquist output indexes provide measures of not only technical efficiency but also of the effects of the changing technology over time/or across production units. To single out the measurement of technical change, note that if the observed input vector X^t and the observed output vector Y^t are on the production frontier, then, consequently, $d^t\left(Y^t, X^t\right)=1$ for $t=0,1$. The above indexes can be restated as follows:

$$\delta^0 = Q_M^0\left(Y^0, Y^1\right) = d^0\left(Y^1, X^0\right) \text{ and}$$
$$\delta^1 = Q_M^1\left(Y^0, Y^1\right) = 1/d^0\left(Y^1, X^0\right),$$

where δ^0 can be interpreted as a measure of the size of Y^1 relative to Y^0 in the context of period 0 technology. Analogously, δ^1 can be interpreted as the size of Y^1 relative to Y^0 in the context of period 1 technology.

Assuming revenue maximising behaviour (cost minimising behaviour) on the part of the production unit for $t = 0, 1$, Caves et al. have shown that the geometric average of the two Malmquist output (input) quantity indexes can be approximated by the Törnqvist output (input) quantity indexes, as defined by the numerator (denominator).

Diewert (1992) defined the Malmquist input quantity indexes, Q0M(X0, X1) and Q1M(X0, X1), in a completely analogous manner. He showed that the Fisher output (input) quantity indexes, as defined in the numerator (denominator), are equal to each of the Malmquist output (input) quantity indexes. Moorsteen (1961) defined the Malmquist TFP indexes as a ratio of the Malmquist output indexes divided by the Malmquist input indexes.

Färe, Grosskopf, Norris, & Zhang (1994) introduced a modification to the Malmquist productivity index suggested by Caves et al. (1982), which requires that (for the output-based measure) firms are revenue maximisers and that (for the input-based measure) firms are cost minimisers. Their calculations exploit the fact that the output distance functions used to construct that Malmquist index are reciprocal to Farrell´s (1957) output-oriented technical efficiency measure. They, therefore, bear a close relationship to the CCR output-oriented DEA model. This link to efficiency allows the decomposition of productivity changes into changes in efficiency and changes in the best-practice frontier (technical change), an idea used by Nishimizu and Page (1982) in a parametric context. They used this model to determine the pattern of hospital productivity in Sweden between 1970 and 1985. By comparing annual changes in the productivity of individual hospitals, they identified both the general trends in productivity of the hospital industry and the individual hospitals exhibiting a pattern of changes in productivity that differ from the rest of the industry.

Färe et al. (1994) analysed the productivity growth in 17 OECD countries over the period 1979 to 1988. The non-parametric programming method was used to calculate the component distance functions of the Malmquist index. The enhanced decomposition model of Färe et al. (1994) was used to decompose the Malmquist productivity change into the components of technical change, pure technical efficiency change, and the change in scale efficiency. This enhanced decomposition takes the efficiency change component calculated relative to the CRS technology and decomposes it into a pure efficiency change component (calculated relative to variable returns to scale [VRS] or VRS technologies) and a residual scale component which captures changes in the deviation between the variable returns and the CRS technology. The results revealed that overall performance in the United States was close to the average for the sample; however, the United States was above average in terms of technical change. The United States consistently shifted the frontier over the entire sample period. Productivity growth in Japan was well above average due, to a large part, to *catching up* to the frontier rather than to technical change (shifts in the frontier).

Ray and Mukherjee (1996) proposed a non-parametric decomposition of the Fisher productivity index into different factors, such as changes in technical and allocative efficiencies, shifts in the cost functions due to technical change, and changes in output attributes. Firm-level data for 21 airlines for the years 1983 and 1984 are used in an empirical application that provides an illustration of the proposed method.

The approach of Ray and Desli (1997) to decompose the Malmquist productivity change into its different components differs from the extended decomposition model proposed by Färe, Grosskopf, and Lovell (1994). The extended model of Färe et al. assumes CRS at the stage of measuring technical change but subsequently switches to VRS to separate the scale effect component, which is not internally consistent. In contrast, the decomposition model developed by Ray and Desli (1997) assumes VRS to measure each and every component of the Malmquist productivity index.

1.5 Parametric Frontier Approach

The empirical literature on frontier technology and the calculation of efficiency measures starts with the path-breaking paper of Farrell, who identified the technical efficiency in terms of realised deviations from the idealised frontier (the isoquant). The approach suggested by Farrell falls naturally into an econometric approach in which the inefficiency is identified with disturbances in a regression model.

Studies on the parametric frontier approach can be classified broadly into two: deterministic and stochastic.

1.5.1 Deterministic Parametric Frontier

Farrell (1957) suggested computing a parametric convex hull of the observed input-output ratios by choosing the Cobb-Douglas production function. Though he acknowledged the undesirability of imposing a restrictive functional form with the idea of being able to express the frontier in a simple mathematical form, he himself did not follow up on his own suggestion. Aigner and Chu (1968), who were first to follow Farrell's idea, suggested a log-linear production function:

$$Q_i = AX_{1i}^{\beta_1} X_{2i}^{\beta_2} U_i = Q_i^* U_i ,$$

where U_i is a random disturbance between 0 and 1. Taking the logarithm of both sides leads to

$$y_i = \alpha + \sum_{k=1}^{K} \beta_k x_{k,i} + \varepsilon_i = \alpha + \sum_{k=1}^{K} \beta_k x_{k,i} - u_i ,$$

where $\alpha = \ln A$, $x_{ki} = \ln X_{ki}$, $\varepsilon_i = \ln U_i$, and $u_i = -\varepsilon_i$.

The non-stochastic part on the right-hand side is viewed as frontier. It is also *deterministic* because the stochastic component of the model is entirely contained in the inefficiency term. Farrell's measure of technical inefficiency is then,

$$Q_i / Q_i^* = U_i = e_i^{-u} .$$

This one-sided error term, labelled as the *inefficiency term*, forces

$$y_i \le \alpha + \sum_{k=1}^{K} \beta_k x_{k,i} .$$

Aigner and Chu (1968) suggested two estimation methods that would constrain the residuals ε_i to be negative:

Linear programming:

$$\theta_{lp} = \min_{\beta} \sum_{i=1}^{n} \left| y_i - \alpha - \sum_{k=1}^{K} \beta_k x_{ki} \right|$$

s.t.

$$\varepsilon_i = y_i - \alpha - \sum_{k=1}^{K} \beta_k x_{ki} \le 0, \forall i$$

and quadratic programming:

$$\theta_{qp} = \min_{\beta} \sum_{i=1}^{n} \left| y_i - \alpha - \sum_{k=1}^{K} \beta_k x_{ki} \right|^2$$

s.t.

$$\varepsilon_i = y_i - \alpha - \sum_{k=1}^{K} \beta_k x_{ki} \le 0, \forall i$$

In the case of the deterministic parametric production frontier, no assumption regarding the distribution of the disturbance term has been made. However, it is essential to make certain assumptions about x and u when the statistical frontier is used for efficiency estimation. The most essential assumptions about x and u are that they are independently and identically distributed (*iid*), and that x is exogenous (independent of *u*). Nonetheless, the possible distribution for u could probably be dependent upon the nature of the structure on the frontier production chosen.

It was Afrait (1972) who first explicitly proposed this model with a two-parameter beta distribution of e^{-u}, which could be estimated through the maximum likelihood estimation (MLE) method. According to Richmond (1974), this amounts to a gamma distribution for u. On the other hand, Schmidt (1976) observed that the Aigner-Chu criteria could be interpreted as the log-likelihood functions for models in which one-sided residuals are distributed as exponential or half-normal. He showed that if u is exponential, then the Aigner-Chu LP model is maximum likelihood, and if u is half-normal, then their quadratic programming model is maximum likelihood.

Nishimizu and Page (1982) developed, for the first time, a methodology that decomposes productivity growth into technological progress and efficiency change for the productivity analysis in the economy of Yugoslavia by specifying a translog production function in the parametric technique of Aigner and Chu (1968) and Timmer (1971). They defined technological progress as the change in the best practice production frontier, and established its rate by direct estimation of a deterministic frontier production function. All other productivity changes – for example, learning by doing, diffusion of new technological knowledge, improved managerial practice, as well as short-run adjustment to shocks external to the enterprise – are regarded as technical efficiency changes.

1.5.2 Stochastic parametric frontier

The deterministic production frontiers discussed so far are based on the idea that all variations in a firm's performance are attributed to the variation in a firm's efficiencies relative to the common family of frontiers shared by all the firms. However, the notion of a deterministic frontier does not take into account the possibility that a firm's performance may be affected by factors entirely outside its control, such as poor machine performance, bad weather, input supply breakdowns, and so on, and by factors under its control labelled *inefficiency*. In effect, the single term *inefficiency,* mixed with the effects of exogenous shocks, measurement error, and inefficiency, is subject to questions. Thus, the picture of the concept of a stochastic frontier emerges from the theories of Aigner, Lovell, and Schmidt (1977), Battese and Corra (1977), and Meeusen and Van den Broeck (1977), who were motivated by the idea that deviations from the production frontier might not be entirely under the control of the DMUs being studied. The idea behind the stochastic frontier is that there may be a measurement error on the dependent variable but not on the independent variables, and that the equation may not be completely specified. Therefore, the error term in the stochastic frontier is composed of two parts: one part (systematic) permits the random variation of the frontier across firms and captures the effect of the measurement error, other statistical noise, and random shocks outside the control of the firm, and the other part (one-sided error term) captures the effect of the inefficiency relative to the stochastic frontier.

Perelman (1995) made use of the stochastic production frontier technique to measure and decompose productivity growth into technological change and technical efficiency change in an international