ALTERNATIVE APPROACHES TO TECHNICAL EFFICIENCY ESTIMATION

ALTERNATIVE APPROACHES TO TECHNICAL EFFICIENCY ESTIMATION: A COMPARATIVE STUDY

By

DADING LI, M.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (1991) McMaster University (Economics)
Hamilton, Ontario

TITLE: Alternative Approaches to Technical Efficiency Estimation: A Comparative Study

AUTHOR: Dading Li, M.A. (NanKai University)

M.A. (McMaster University)

SUPERVISORS: Professor A. Leslie Robb (Chairman)
Professor M.W. Luke Chan
Professor Dean Mountain

NUMBER OF PAGES: xii, 186.

ABSTRACT

Pursuing efficiency is a fundamental characteristic of economic activity. Correspondingly, efficiency measurement seems an eternal interest of production economists. The present dissertation is a comparative study of alternative technical efficiency estimation methods. Two recently developed methods based on different methodologies, namely, data envelopment analysis (DEA) and the stochastic frontier approach (SF) are studied. In this dissertation we review the production and efficiency structure defined by modern production theory. Based on earlier works of Afriat, we discuss a set of propositions underpinning the non-parametric programming approach (or DEA). Further, we demonstrate the relationship between non-parametric and parametric production frontiers as references for technical efficiency measurement. We also explore the corresponding relationships between various versions of the DEA model and their implications regarding returns to scale properties. On the side of the SF approach, we work out a conditional estimation model to extract technical efficiency from a composite error structure. The main empirical contribution is a simulation study that is carried out to examine the capabilities of both approaches under various circumstances. In the first set of experiments we examine the performances of the two methods under assorted efficiency profiles, by which we describe the industry's efficiency distribution. Then, in a second set of experiments we investigate the performance of the two methods when the experimental data has different returns to scale properties. Finally, we test the robustness of the two models in regards to varied magnitudes of random noise. Our results indicate that though the SF model often leads the competition by a small margin in our experimental environment, both methods have reasonable performances.

ACKNOWLEDGEMENTS

At the time to submit this final fulfilment for the requirement of my degree study, I feel deeply that without help and support from many people, my own efforts could not have resulted in the present dissertation .

First, I would like to acknowledge and thank the members of my dissertation committee for their guidance, patience and support. In the last few years, Professor Les Robb, Chairman of the Committee, guided my research by numerous thoughtful and stimulating advices extending from the arrangement of the entire investigation scheme to the details of presentation. Professor M.W. Luke Chan, who has given me much generous help since I arrived in this country, and Professor Dean Mountain have both given many precious suggestions and insightful comments by using their expertise in this research area. The knowledge and research experience accumulated under their guidance will be a priceless asset in my future endeavour.

An additional thank goes to the Canada International Development Agency (CIDA). Its initial sponsorship enable many Chinese students, including me, to pursue their studies in this country.

Finally, my thanks should be dedicated to my late father, my mother and in particular my wife, Hong. Without their support and encouragement, I would not have been able to finish my studies. Moreover, I am in a great debt to Xiaoli, my son, for his patience and well-behaved cooperation during my working time.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure

LIST OF TABLES

Table

CHAPTER 1. INTRODUCTION

The present dissertation is a comparative study of alternative technical efficiency measurement techniques. The importance of the subject can not be overstated in modern economics. As Michael J. Farrell, one of the pathfinders in this area, expressed, "the measuring of the productive efficiency of an industry is important to both the economic theorist and the economic policy maker".[1]

The theoretical structure of productive efficiency analysis was laid out in early nineteen fifties by T. C. Koopmans (1951), G. Debreu (1953) and M.J. Farrell (1957). However, more comprehensive studies and research have emerged only in the last ten or fifteen years. It has taken a few decades for the profession to respond to the earlier thrusts. A couple of reasons may account for the delayed concern. First, the primary interest of empirical production analysis had been focused mainly on the functional relationship in the production process. Those functional relationships are characterized by the economic behaviour of production agencies at full efficiency. Until those relationships were fully

---

[1] M.J. Farrell (1957), 11.

exposed and understood, the analysis of productive inefficiency, which primarily concerns non-optimal or sub-optimal status, had only secondary importance. Second, the empirical studies of productive efficiency rely upon support from other branches of the science (e.g. applied mathematics, statistics and operation research) and hinge on more efficien computing techniques and facilities. These supports were either in a less developed stage, or unavailable, in earlier years. When the first linear production frontier was modeled by Farrell with the assistance of EDSAC (Electronic Delay Storage Automatic Calculator) in the middle fifties, few people had access to equivalent facilities. Recently, with increasing attention and updated technical assistance, studies of productive efficiency have developed in many dimensions. To appreciate the recent achievements, it is worthwhile to revisit the historical path of development over the years.

The study of productive efficiency was primarily a post-World War II phenomenon. The 1950's was the period when the foundations of productive efficiency studies were laid out. The newborn set-theoretic model, and the related activity analysis in particular, were the main thrusts to initiate the analysis of productive efficiency. The following ideas from activity analysis had an immediate influence on the efficiency studies: (a) the production frontier of an industry can be defined by a group of firms with best performance; (b) a firm's production potential can be defined as the linear combinations of the activities which have best performance; (c) the efficiency of a firm is measured by its position relative to the frontier. Though efficiency studies have changed in many ways in these days, the thoughts are still the generally valid principles guiding the current analysis.

An important theoretical advance in 1950's was the structure of productive efficiency proposed by G. Debreu (1951). Though the structure was developed for a general equilibrium framework, it has a direct impact on production theory. According to Debreu, productive inefficiency may be decomposed into two parts: allocative (in)efficiency and technical (in)efficiency. While the former measures the achievement of an economy in choosing the optimum input bundle given the factor prices, the latter assesses the success of an economy in producing maximum output from a given inputs bundle. Concerning technical efficiency measurement in particular, Debreu introduced a "coefficient of resource utilization", which measures the radial distance between an interior point and a corresponding boundary point of a production possibility set.

Based on these theoretical concepts, the first workable measurement scheme emerged in 1957 when a pathbreaking paper by Michael J. Farrell appeared. Farrell's first contribution is his decomposition of productive efficiency which is a analogue of Debreu's work but in the framework of production analysis. Moreover, while the analytical structure embraced Debreu's ideas closely, his measurement procedure was deeply influenced by activity analysis. According to Farrell, the technical efficiency of an observation can be measured by the radial distance between the point and the corresponding point on a production frontier, which is built from the data through a linear programming procedure.

Farrell's work has some important features. First, the production frontier, which is constructed from "best performance" input-output data, is consistent with the theoretical definition of a production function, i.e. the maximized output for given inputs. Second, the linear programming procedure can be extended to cover multi-input, multi-output situations without any difficulties. It is well known that this type of production process usually can not be handle properly without resorting to an indirect method such as estimating the cost function. Third, the technical inefficiency of an economic agency can be revealed solely based on the analysis of the physical input-output data. As a result, efficiency analysis becomes possible for non-profit organizations, public sector, and even agents in a non-market driven economy where the market information is unavailable.

Two major criticisms to Farrell's measurement approach are frequently encountered:[2] one regards its inability of handling random errors and the other concerns the lack of flexibility in handling various degrees of returns to scale in the production process. Farrell and Fieldhouse (1962) attempted to remedy the returns to scale problem but without much success. Developments since then indicate that the solution to the problem depends on compatibility between the measurement and a reasonably regulated production structure. With a properly defined production structure, the Farrell measurement can be applied to the production process with constant returns to scale, non-increasing returns to scale and variable returns to scale property.[3] The other criticism seems to be a more fundamental problem. Being non-statistical, Farrell's measurement is inherently unable to accomodate random errors. However, the validity of the measurement may depend on whether the random noise plays a dominant role in a production process. Farrell's measurement still be valid as long as random error is ignorable comparing to other systematic factors.

The remainder of the sixties seems to be a less active period for the studies of productive efficiency. The work by Aigner and Chu (1968), however, is a major thrust in the

---

[2] See Royal Statistical Society (1957).

[3] See the discussion in Chapter 4.

period. It is well known that programming approach features a non-parametric frontier property. In other words, the production frontier is defined relatively rather than parametrically in the approach. However, Aigner and Chu attempted to impose a parametric form to the programming approach. Briefly, the technique attempted to build up and locate a parametric production function by minimizing the radial distance between all the data points and a theoretical frontier yet to be formed. Correspondingly, the programming problem need not have a linear form. The production function derived by Aigner and Chu has a frontier property, hence is consistent with the theoretical definition of a production function. Further, all the neo-classical properties are preserved. Since the production frontier has an efficiency property, the deviation of an observation from the frontier must be interpreted as inefficiency. This treatment positively confirms the presence of technical inefficiency in a model of production process. On the other hand, because the frontier is still obtained through a programming method, it shares some fundamental weaknesses with Farrell's procedure, e.g. there is no room for a stochastic component in the model. Aigner and Chu's approach, however, had a theoretical impact reaching beyond its modelling techniques. The concept of a "frontier production function" was introduced explicitly in production theory. The traditional treatment of production function estimation was seriously questioned and a profound discussion on the relevant issues was stimulated in the subsequent period.[4]

Studies of productive efficiency were not truly activated until Afriat's paper (1972) emerged. Afriat's work bestowed efficiency studies with two major contributions. The first contribution involves a series of theorems which shows that any data dependent production frontier is bounded from below by a data dependent linear production frontier.[5] This linear frontier is in fact an inner envelope or "underwear" covering the data set. Therefore, any efficiency measurement taking some production frontier as a reference must be bounded by the efficiency measurement that takes the "inner" envelope as the reference. These theorems provide the different approaches, e.g., conventional functional approach and the programming approach, with a common shell and a mutually accepted efficiency structure.[6]

---

[4] See D. J Aigner and S.F. Chu (1968), P. Schmidt (1976) (1977) and S.F. Chu (1977).

[5] A data dependent production frontier, roughly speaking, is a production function spanned by best performance firms and covering the entire data set. It, hence, has a frontier property. See Chapter 4 for a detailed discussion.

[6] In the Data Envelopment Analysis (DEA), a programming approach developed later on, the root was traced back to the inverted cost function introduced by Shephard (1953). However, one may note that the DEA formulation has a closer connection to Afriat's Theorems. See R. Banker, A. Charnes and W. W. Cooper (1984).

Second contribution by Afriat is the concept of "the distribution of technical inefficiency", which describes how the efficiency indexes or scores are distributed among the firms within an industry. A frequency distribution was initially introduced into efficiency analysis by Farrell to summarize the results of estimation.[7] The distribution, nevertheless, was adopted in a somewhat ad hoc way and it was by no means an integrated part of Farrell's theoretical efficiency structure. Afriat put the hypothesis in a more formal way: the data points laid inside the production frontier could be modeled by some sort of statistical density function. As he demonstrated, a Beta distribution might serve the purpose adequately. The significance of the distribution can be stretched far beyond Afriat's original intent. The distribution may, in fact, reflect an industry's dynamic profile revealing technical progress, technology transformation and diffusion, maturity, etc.. It is of considerable interest to note that the concept of "capacity distribution" by Johansen (1972) contains a similar idea. This distribution is a snapshot of the productivity profile for an industry and bears the same efficiency interpretation (according to Johansen). The coincidence suggests nothing but an emerging methodology which takes the efficiency profile as

[7] See M. J. Farrell (1957), 270-271.

a vehicle to address an industry's structure.

The earlier theoretical developments reaped a considerable harvest in the seventies. Many efficiency estimation techniques were cultivated in this period. Traditional econometrics brought about two new procedures in the field of production function estimation. A so-called "deterministic statistical model" estimates the frontier production functions by using information of an efficiency distribution. In this approach, the frontier property of a production function is emphasised and technical inefficiency is treated as a statistical variable spreading beneath the frontier. The second approach suggested a stochastic production frontier which asserts that the production frontier is a distribution rather than a fixed measure. The stochastic frontier model was initially introduced by Aigner, Lovell and Schmidt (1977) and has been applied by a larger group of researchers.[8]

Totally different from the econometric treatment, another important development in late seventies is the so-called "data envelopment analysis" (DEA) which emerged from the area of management science. DEA efficiency measurement, formulated by A. Charnes, W. W. Cooper and E. Rhodes (1978),

---

[8] Most referenced works in the group are W. Meeusen and J. van den Broeck (1977), F.R. Forsund and Hjalmarsson (1979), F.R. Forsund and E.S. Jansen (1977), R.J. Kopp and V.K. Smith (1978) and L.F. Lee and W.G. Tyler (1978).

is an alternative version of Farrell's efficiency measurement. However, the measurement is endowed with a brand new formulation which bears the colours of operation research.

The early 1980's witnessed a surging interest in frontier production function and in efficiency studies from the economics profession. As an extension of the earlier stochastic frontier model, a new conditional estimation technique was introduced to extract a measure of technical efficiency from a composite error structure. The method, suggested by J. Jondrow, C.A. Lovell, I.S. Materov and P.Schmidt (1982), defined a new standard for technical efficiency estimation. At the same time, the DEA approach was also extensively studied from different angles. In addition to the original constant returns to scale formulation, some new formulations were introduced to cover the production process characterized with non-increasing returns to scale and variable returns to scale [see Banker, Charnes and Cooper (1984)]. Assisted by better estimation techniques, DEA has now been applied to efficiency and productivity investigations in many areas. These include: manufacturing industry, agriculture, administration, education, airlines, energy, etc..[9]

Many innovative thrusts have emerged in the most

---

[9] See Lovell and Schmidt (1988) for more references.

recent period. The stochastic frontier model and the DEA model have became dominant approaches to productive efficiency analysis. On the side of the stochastic frontier approach, many efforts have been spent to extend the analysis to the cost function and related systems.[10] Another focus of the approach is the search for more flexible efficiency structures. Greene (1990) proposed a Gamma-distributed efficiency structure and has worked out the formula for the conditional estimation of technical efficiency. There are also some new developments in the estimation techniques for the stochastic frontier model. The "moments" method has drawn considerable attentions and has been applied in various circumstance [see Kopp and Mullahy (1990), Greene (1990)]. Moreover, Kumbhakar (1987, 1989, 1990) has applied the stochastic frontier model to panel data. As a result, technical efficiency has become a time-varying factor. One may find that there is a strong dynamic potential in the stochastic frontier analysis along this orientation.

On the side of the DEA approach, in addition to the sprouting of empirical applications, a major development is the comprehensive analysis of the cone structure of the model. According to this analysis, the frontier structure can be adjusted by imposing a priori weight vector (multipliers). As

---

[10] For a more detailed survey of recent development in the stochastic frontier approach, see Bauer (1990).

a result, flexibility and manoeuvrability of the DEA model are considerably enhanced [see Charnes, Cooper, Huang and Sun (1990) and Thompson, Langemeier, Lee, Lee and Thrall (1990)]. Further, there have been recent efforts to introduce stochastic components into the programming model [Sengupta(1990)]. In the empirical literature, we also find comparative studies which assess the stochastic frontier and the DEA model by putting them into a competition in empirical work [e.g., Bjurek, Hjalmarsson and Forsund (1990)].

As a result of these dynamic developments, the analysis and measurement of productive efficiency has been integrated gradually into modern production theory.[11] These developments provides a primal motivation for present comparative study.

Although allocative efficiency, a part of productive efficiency, is a primary concern in production theory for many reasons, the domain of the present study is confined to technical efficiency estimation for the following reasons. First, knowledge and information about total productive efficiency is based on the understanding of its two complementary components. Should technical efficiency be well understood and measured, allocative efficiency can be estimated without much additional complication. To some

---

[11] For theoretical generalization of the developments, see R. Fare (1988) and Varian (1984, 1990).

extent, the technical efficiency is a starting point for any further investigation of production process. Second, technical efficiency measurement has its own economic significance which has often been ignored. Economists are used to thinking of measurement of technical efficiency as a job for engineers. This is simply not the case. An economist's measurement of technical efficiency never relies on any predetermined physical or technological standard as a yardstick. Technical efficiency measured by economists is comparative efficiency with a strong Pareto sense. This has been expressed fairly clearly in both the earlier literature of activity analysis and in more recent literature [See Koopmans (1951) and Charnes, Cooper and Rhodes (1985)]. Third, technical efficiency measurement has much potential in an economy or a section of an economy which is not market driven or where price information is not available. In these circumstances, technical efficiency measurement is the only possibility.

This introduction can be concluded as follows: (i) Debreu-Farrell's efficiency structure is a corner stone of the productive efficiency analysis shared by different approaches. (ii) The current research frontier is led by two key approaches: the stochastic frontier production function approach and the nonparametric programming approach (or DEA, as named by management scientists). (iii) Though both approaches have been extensively applied, comparative studies seemed to be an rarely explored area. All these observations merge to a major driving force to pursue the present study: a comparative study of two alternative approaches to technical efficiency measurement.

The thesis is structured as follows. Chapter 2 reviews the production structure and the notion of efficient technology in modern production theory. Chapter 3 examines various technical efficiency estimation models and explores methodological aspects of technical efficiency estimation. Chapter 4 discusses empirical frontier production functions and introduces a set of "inner envelope" propositions which prove that the family of empirical frontier production functions are bounded by a set of linear production frontier. Based on these propositions, we focus on the programming approach (the DEA model). Various model specifications and their implications are discussed. Chapter 5 is devoted to the stochastic frontier approach. Based on previous studies, we work out conditional estimation procedures for a generalized stochastic frontier model and discuss its implications. Chapter 6 lays out the ground work for a simulation study of the two approaches. Experimental design and relevant test statistics are the topics of the chapter. In Chapter 7, we report the results of the experimental simulation. The dissertation is summarized in Chapter 8.

CHAPTER 2 THE PRODUCTION STRUCTURE AND EFFICIENT TECHNOLOGY

This chapter concentrates on the theoretical background of this productive and technical efficiency study. Section 1 of the chapter reviews a set of axioms that regulate a production process. The primary function of the axiomatic approach is to set out a well regulated production structure. Based on this structure, Section 2 examines various definitions of an efficient production technology. With a well defined theoretical structure, empirical production behaviour can be reasonably interpreted. Section 3 concentrates on the measurement of productive efficiency. Section 4 concludes the chapter.

2.1 Basic Axioms of the Production Correspondence

Our choice of axioms is based on three major considerations: first, an axiom must be able to interpret production data with reasonable logic; second, an axiom has to be flexible enough in a sense of "simple and true" so it will not lose validity if applied to various production will not lose validity if applied to various production models; finally, it should be strong enough to have an active regulating role and yield some useful results.

The set of axioms is aimed to regulate a production process described by a mapping from output space R^m_+ to input space R^n_+ (an input correspondence):[1]

$$L: R_+^m \to R_+^n$$

By definition, **L** is a "point to set" mapping. In other words, the image of an output **u** in the input space is the input set **x** that can produce at least **u**. Correspondingly, the graph of **L** is defined as:

$$GrL(u) = [(x,u) \mid x \in L(u),\ u \in R_+^m] \quad \textbf{(2.1)}$$

The input correspondence **L(u)** is regulated by the axioms stated as follows:

L.1 L: $R_+^m \to R_+^n$ is a closed correspondence.

The axiom states that the graph of the input correspondence is a closed set. By definition, a set is closed if all its boundary points exist and are contained in the set. In Figure 2.1, the boundary set for graph L is the

---

[1] For the sake of simplicity, our attention is limited to the input space unless otherwise stated. For a detailed discussion on the output correspondence and complete axioms, see R. Fare (1988).

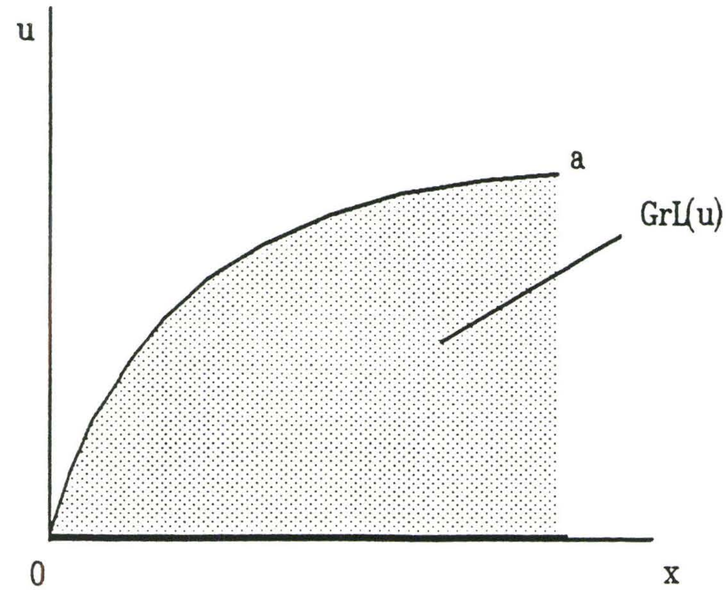

Figure 2.1 The Graph GrL(u)

set of the frontier points from **o** to a.

The closedness axiom ensures the existence of a production frontier (or isoquant in input space) and continuity of the frontier. As we will explain later, efficient technology is a sub-set of the frontier. Thus the closedness further ensures the existence of an efficient production set.

L.2 if $u \geq 0$, *then* $o \notin L(u)$

This axiom states that if at least one element of the output vector **u** is larger than zero, then its image, **L(u)**, cannot have all zero elements. In other words, a production process cannot produce something from nothing.[2] Figure 2.2 illustrates the axiom graphically: if the axiom is true, the vertical axis cannot be part of the graph of the input correspondence except at the origin.

$L.3 \quad L(0) = R_+^n$

The third axiom states that a positive input may result in a null output, so the production activity may turn out to be a fruitless effort. To put it in another way, there is a possibility of extreme inefficiency. In single input, single output space, the axiom implies that the graph of the production function may collapse onto the horizontal axis as shown in Figure 2.2.

$L.4 \quad \text{for all } x^* > x \in L(u), \quad x^* \in L(u)$

This axiom states that if $\mathbf{x}$ belongs to the set of inputs which is able to produce $\mathbf{u}$, then $\mathbf{x}^* > \mathbf{x}$ (each element of the former vector is larger than the corresponding element of the later) must be able to produce $\mathbf{u}$ also.

The axiom has some implications somewhat more complicated than its superficially simple appearance. First,

[2] The axiom comes from a fundamental postulate proposed by Koopmans: "the impossibility of the land of Cockaigne". See Nikaido (1970), 216.

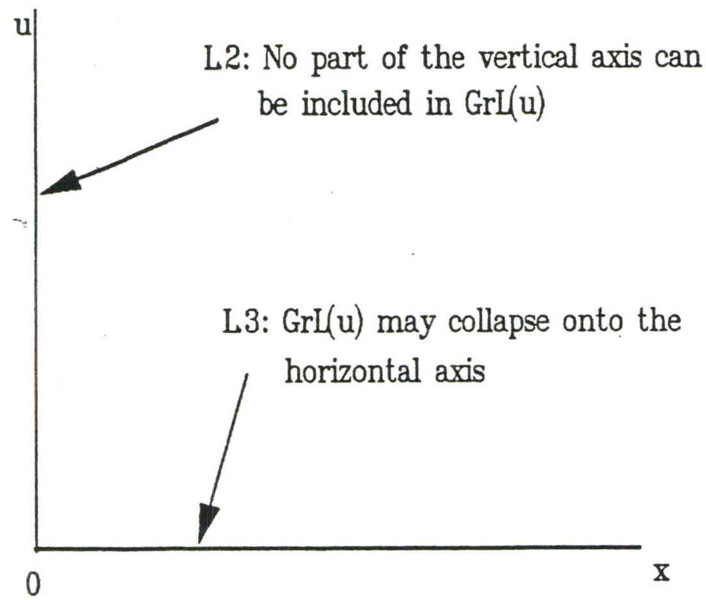

Figure 2.2 The implications of axioms L2 and L3

the production frontier (the boundary of the graph **GrL**) can not decrease. In Figure 2.3.(a)., **x** is an input level producing **u**. The axiom requires that any $\mathbf{x}^* > \mathbf{x}$ must be able to produce **u** also. Then the graph of the correspondence must at least include the area between **o-a-b** and X-axis. As a result, the production frontier (boundary point of the graph, i.e., **o-a-b**) is non-decreasing. The economic interpretation of the result is: we cannot have a situation where all inputs can be reduced with an increasing output level.

Further, a backward bending isoquant in input space is also prohibited by the axiom. In Figure 2.3.(b), **x** is a vector which is able to produce **u**. Suppose there is $\mathbf{x}^* > \mathbf{x}$, the axiom in fact states: (a) $\mathbf{x}^*$ must also be able to produce **u**;

(b) if **u** is produced by x^*, the activity (**u**, $\mathbf{x}^*$) does not attain its potential. The second part of the statement can be verified by a contradictory situation. Suppose the activity ($\mathbf{u},\mathbf{x}^*$) is at its potential, then the isoquant will have a backward bending portion $\mathbf{x}-\mathbf{x}^*$. If this were true, any other point like $\mathbf{x}' > \mathbf{x}$ would have an output level less than u. This is contradictory to the axiom. Therefore, any backward bending portion of isoquant is incompatible with the axiom.

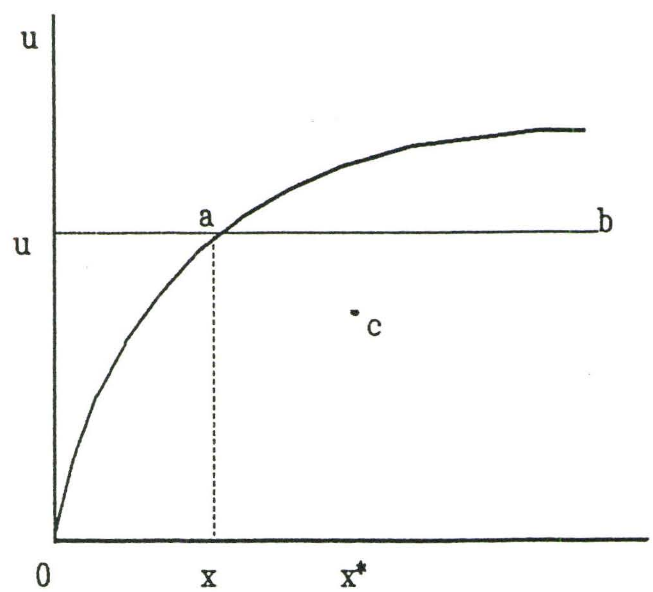

Figure 2.3.(a) The production frontier is non-decreasing in X

The axiom has an important implication in the context of efficiency study. Since the production function is non-decreasing and a backward bending isoquant is ruled out, the behaviour of points like **c** in Figure 2.3.(a)or $\mathbf{x}^*$ and $\mathbf{x}'$ in Figure 2.3.(b)have to be interpreted as a sort of inefficiency.

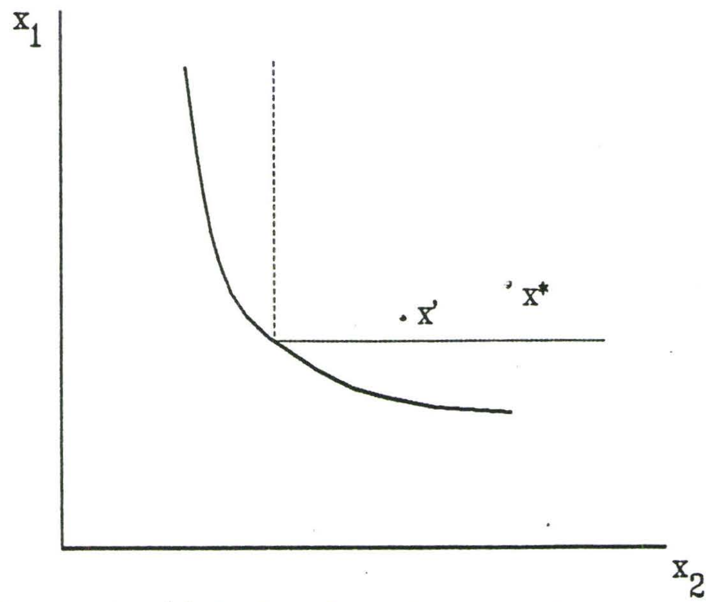

Figure 2.3.(b): Backward bending of an isoquant is prohibited

L.5 If $u^* < u$, then $L(u) \subset L(u^*)$

This axiom, though pertaining to an input correspondence, can be illustrated more clearly in output space. By the axiom, an output vector $\mathbf{u}^* < \mathbf{u}$ (every element of $\mathbf{u}^*$ is less then the corresponding elements of $\mathbf{u}$) should have a larger input image set $L(u^*)$ (Figure 2.4.(a)). Therefore, the production possibility frontier cannot have an inward bending portion (Figure 2.4.(a)).

The possibilities of inefficiency exist according to this axiom. Any output vector less than **u** must be able to be produced by less inputs (thus input set is larger than **L(u)**), if not, inefficiency occurs. Therefore, the production frontier cannot bend backwards toward the u-axis.

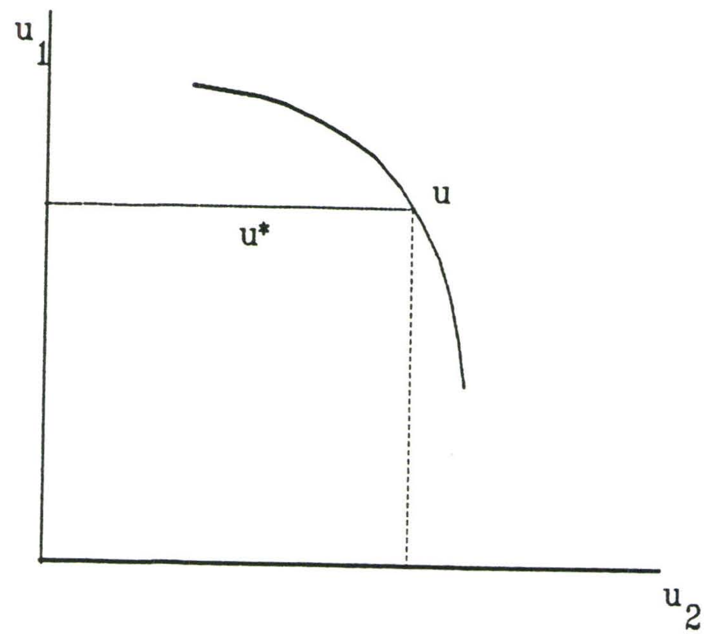

Figure 2.4.(a) The production possibility frontier should not have an inward bending portion

L.6 L(u) is bounded for u < +∞.

The axiom states that for any finite output, the input requirement must be bounded by some finite number. Therefore it rules out the situation where a finite output level requires infinitely large inputs. To extend this axiom to output space, one would conclude that it is impossible to produce infinite output by finite inputs.

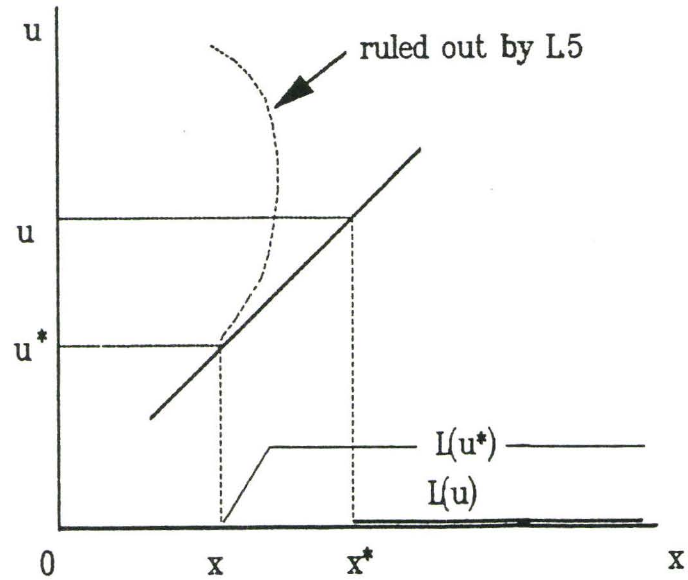

Figure 2.4.(b) The production frontier should not have a backward bending protion

L.7 L is convex on $R_+^m x R_+^n$.

The axiom states that the graph **GrL** is a convex set on $R_+^n x R_+^m$. This implies that the production possibility set **L(u)** is convex too. In Figure 2.5.(a), the graph **GrL** is a convex set and the projected image **L(u)** must have a convexity property.

However, strict convexity of GrL is not required. As a result, the production frontier could have either linear or non-linear structure. [Figure 2.5.(b)]

The above axioms can be matched by a set of symmetrical counterparts regulating the output correspondence.[3]

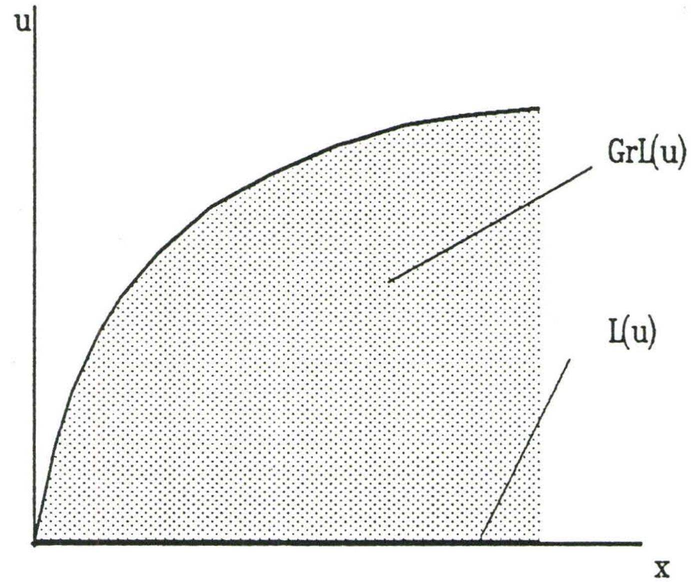

Figure 2.5.(a). Convexity of GrL(u)

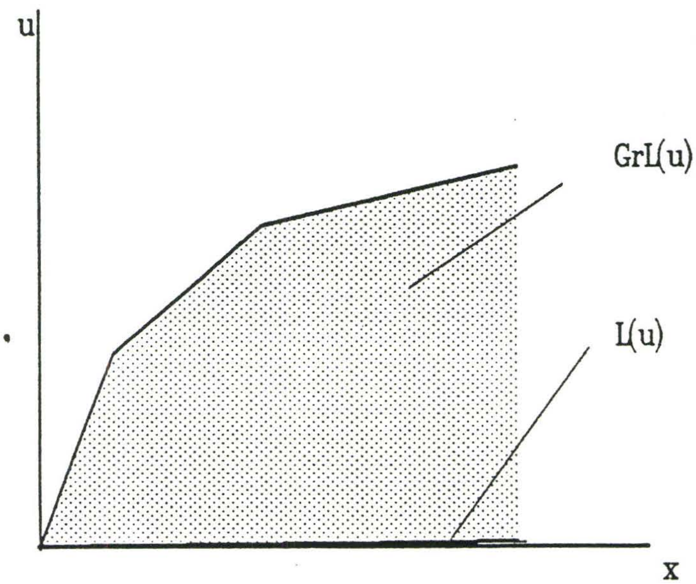

Figure 2.5.(b) Convexity of GrL: a linear boundary

---

[3] See R. Fare (1978), Chapter 1.

In the next section, we formally define an efficient technology and will frequently refer back to the axioms to check their implications for technical efficiency analysis.

2.2 The Efficient Technology

The axioms **L.1** (the closedness of the graph **GrL**) enables us to define the boundary of **GrL.** For an output level **u**, the projection of the boundary of **GrL(u)** on input space can be defined as an isoquant:

$$ISOQ\ L(u) = \{\ x :\ \ x \in L(u),\ x \notin L(u'),\ \ u' \geq u\ \}$$

However, efficiency is not an unambiguous property for the isoquant unless axiom **L.4** is in effect. Axiom L.4 rules out the case that an isoquant has a backward bending portion. Thus, a Weakly Efficient Set (a subset of ISOQ L(u))can be defined as:[4]

$$WEFF\ L(u) = \{\ x :\ \ x \in L(u),\ x^{*} \notin L(u),\ \ x^{*} > x\ \}$$

---

[4] In the following definitions, " **a** > **b** " implies that each elements of **a** is larger than **b** and " **a** ≥ **b** " implies that at least one element of **a** is larger than **b.**

Further, at times we might wish to rule out the case that an isoquant has a portion parallel to the vertical or the horizontal axis. In this case, an Efficient Set can be defined as

$$EFF\ L(u) = \{\, x : \ x \in L(u),\ x^{*} \notin L(u),\ \ x^{*} \geq x \,\}$$

There are nested relationships among the above three input sets as follows:[5]

$$\mathbf{EFF\ L(u)} \subseteq \mathbf{WEFF\ L(u)} \subseteq \mathbf{ISOQ\ L(u)}$$

In Figure 2.6, the curve **a-b-c-d** is an isoquant for **L(u)**. On this isoquant, the backward bending section **a-b** is an inefficient part; the portion **b-c** is weakly efficient and the portion **c-d** is efficient. Therefore, an input vector **x** is technically efficient for an output vector **u** if and only if **x** $\subseteq$ **EFF L(u)**.

Since most currently available efficiency measurement techniques are not able to distinguish between weak and ordinary efficiency, we will not stress the distinction between the two in the rest of the presentation unless otherwise stated.

From the definition of technical efficiency and axiom **L.7**, we define an efficiency subset of the graph:

$$EFF\ GrL = [\,(x,u) \mid x \in EFF\ L(u),\ u \in R_{+}^{m}\,]$$

---

[5] $\subseteq$ indicates a subset.

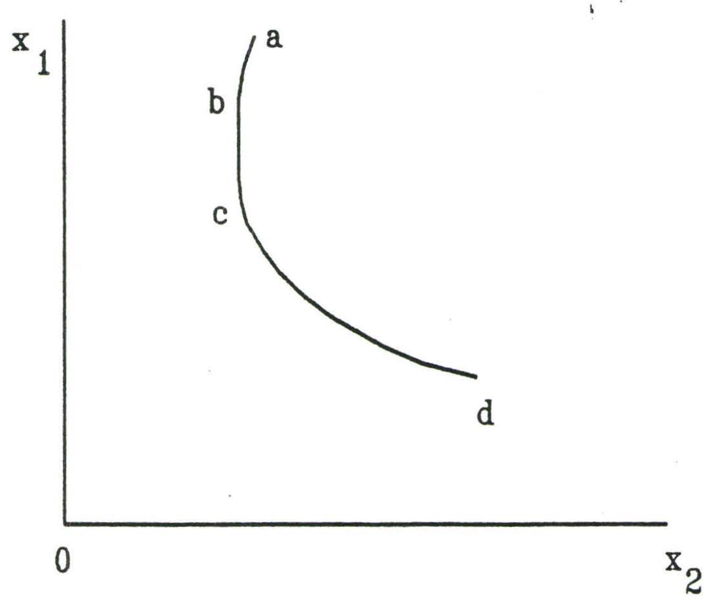

Figure 2.6 The isoquant and the efficient frontier

Note that since **GrL(u)** is a convex set by **L.7, EFF GrL** can be viewed as a function concave on **X**. Accordingly, we may define an efficient production technology as a one-to-one mapping from the input to the output space:

$$f\colon x \to u$$

which has a graph identical with **EFF GrL.**

2.3 The Structure and Measurement of Productive Efficiency

Taking the efficient technology, **EFF L(u),** as a reference set, the technical inefficiency of an interior point

x' in **L(u)** can be measured by it's radial distance to the efficient frontier. The Farrell measurement of technical efficiency is given by

$$T(x' \; ; \; u) = \min \{ \tau : \tau \, x' \in L(u) \; , \tau \geq 0 \} \tag{2.2}$$

where τ is non-negative scaler. Suppose $\mathbf{x}^a \in$ **EFF L(u)** is a referenced point, Farrell's measurement can be made in such a way that $\mathbf{x}^a = \mathbf{T(x'; u)*x}$. Since $\mathbf{x} > \mathbf{x}^a \in$ **EFF L(u)**, **T(x'; u)** must have an effective range **(0,1]**.[6]

Although we have indicated earlier that we will restrict our attention to the technical efficiency, it is useful to note how this relates to overall efficiency. The overall performance of an economic agency can be measured by the productive or total cost efficiency. Given an input price vector $\mathbf{w} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$, the total cost function $\mathbf{C(u;w)} = \mathbf{wx}^*\mathbf{(u;w)}$ reflects the cost for the optimal choice of input set $\mathbf{x}^*$. Thus the total cost or productive inefficiency for an observation **x** is measured by

$$P(\, x \; ; \; u \; , w) = \frac{C(\, u \; ; \; w\,)}{w \, x} \tag{2.3}$$

---

[6] By axiom L.2, it is impossible to produce something from nothing, therefore, T(x'; u) can't be zero for an output vector with at least one non-zero element.

The allocative efficiency can then be defined as the residual part of total productive efficiency after the effect of technical efficiency is removed.

$$A(x; u) = \frac{P(x; u)}{T(x; u)} \tag{2.4}$$

In Figure 2.7, productive inefficiency of **e** producing u is measured by the ratio of the radial distance **oe^3/oe.** Technical inefficiency is measured by **oe^2/oe** and allocative inefficiency by **oe^3/oe^2**. Therefore:

Productive Inefficiency = Allocative inefficiency X Technical inefficiency.

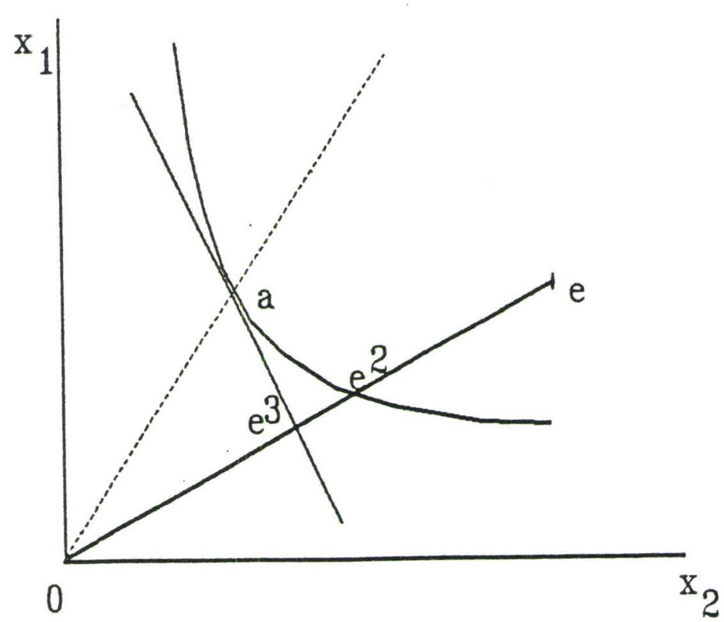

Figure 2.7 Decomposition of total productive inefficiency

Although allocative efficiency and cost efficiency are very important aspects of efficiency analysis, we have no intention to go beyond the scope of technical efficiency measurement in the present dissertation due to the reasons stated in Chapter 1. Moreover, Farrell had revealed the complementary relationship between the technical and the allocative efficiency. If the technical efficiency is known and market information is available, the allocative efficiency can be found without much difficulty.

One may note that Farrell's measurement is rigorously rooted in the modern production theory. The referenced technology in Farrell's measurement is exactly the efficient production frontier as we defined in the previous section. The measurement itself is a perfect match of Debreu's "coefficient of resource utilization". Moreover, as noted by Lovell and Schmidt (1988), the measurement **P(x;u)** has a duality relationship with **L(u)** and in fact is the inverse of the distance function proposed by Shephard (1953,1970).

However, if the referenced production frontier is weakly efficient, in particular, a vertical or horizontal segment of an isoquant, for example, the Farrell measurement of technical inefficiency may yield a biased measurement. In Figure 2.8, the referenced frontier **b-c** is a weakly efficient portion, so the radial measurement is obviously an overstated measuring. The Lieontief technology is an extreme example. In

Figure 2.8, the line **b-c-d** is a Lieontief production frontier. The efficient frontier then collapses to a point **c**. Obviously, any radial estimate not passing through the point would overestimate the true efficiency. However, this flaw is shared by all the radial efficiency measurements.

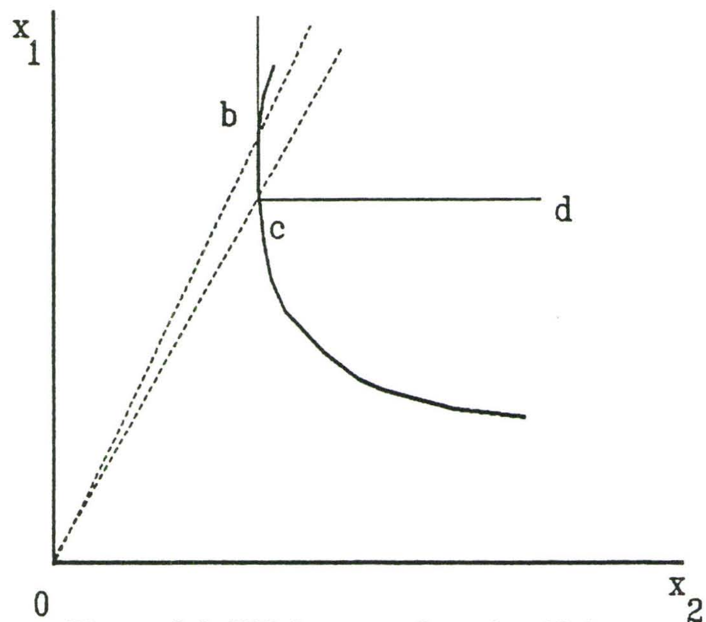

Figure 2.8 Efficiency and weak efficiency

To correct the weakness, Fare and Lovell (1978) suggested a Russell measure of technical inefficiency which gauges the non-radial distance of technical inefficiency as:

$$R(x\ ;\ u) = \min\left\{\frac{\sum_i \delta(x_i)\ \lambda_i}{\sum_i \delta(x_i)}\ ;\ \lambda x \in L(u)\ ,\ \lambda_i \in (0,1]\right\} \quad (2.5)$$

where

$$\delta(x_i) = \begin{cases} 1 & , \ if\ x_i > 0. \\ 0 & , \ otherwise. \end{cases}$$

$$\lambda = (\lambda_1, \ldots, \lambda_n)$$

Note that the reference here for a data point is not taken along the radial direction but along each dimension of the input space. Correspondingly, the original efficiency measuring is an n-dimensional vector that can be converted into a weighted scaler. Apparently, the Farrell's measurement is a spacial case of the Russell's measurement as $\lambda_1=\lambda_2=\ldots,=\lambda_n$.

However, there are a few problems with Russell's measurement. First, the technical inefficiency gauged along the different directions may not maintain a measuring consistency. Second, the local property and behaviour of the referenced frontier play a more important role than in the Farrell's measurement. Unless the knowledge about the referenced frontier is complete, Russell efficiency should be interpreted with caution. Moreover, estimation of Russell efficiency is not as tractable as Farrell's. Therefore, Farrell's measurement is still believed to be the most tractable measure of productive and technical efficiency and we will use it in this thesis.

2.6 Summary

In this chapter, we have reviewed the theoretical background for efficiency measurement. The fundamental axioms of the modern production theory are the starting point in defining the structure of production and efficiency. Unlike the calculus based classical framework, in this structure, we do not assume that economic agencies operate in a situation where all marginal conditions being fulfilled. On the contrary, inefficient production and slackness in operation is plausible. In this chapter, we also reviewed Farrell's decomposition of productive efficiency and the different concepts of efficiency. In the next chapter, we turn to empirical issues: the estimation techniques for technical efficiency. We will review various modelling techniques and discuss the technical efficiency from a different point of view.

CHAPTER 3. EMPIRICAL MODELS OF TECHNICAL EFFICIENCY MEASUREMENT

The theoretical structure of technical inefficiency discussed in the Chapter 2 provides a framework which shelters various production analyses. Nevertheless, it is well known that in practice "... there is not yet a consensus on how one should, or whether one can, measure the technical efficiency of a firm, even if this agreed to be useful things to measure."[1]

This chapter, presented in five sections, reviews alternative modelling techniques and their underpinnings.[2] The first section reviews efficiency measuring techniques in early classical production analysis. The production frontier models developed in 1970's are discussed in section two. In section three, we concentrate on two more updated approaches: the stochastic frontier model and the non-parametric programming model. The recent development of efficiency

---

[1] F.R. Forsund, C.A. Lovell and P. Schmidt (1980) 23.

[2] For the reviews of efficiency studies in different period, see Nerlove (1965), Kopp (1981), Silkman (1986), Sengupta (1988), Lovell and Schmidt (1988), Bauer (1990) and Seiford and Thrall (1990).

estimation techniques is described in section four. In section five, we discuss the significance and implications of the recent developments. The last section summarizes the chapter.

3.1 Efficiency Measurement in the Classical Production Model

The notion of technical inefficiency seemed to be an ambiguous concept in early production analysis. Technical efficiency was often viewed as a pure technical measuring of the production process rather than a relative economic measuring. The economic significance of technical efficiency measuring was not fully recognized.[3] However, some classical analysts did provide penetrating insights on technical efficiency. In an influential pioneering work of production theory, Marschak and Andrews (1944) stated[4]

> *... the production function will change, even within the same industry, from*

---

[3] For example, the constant term in a Cobb-Douglas production function was interpreted as a "term of technical efficiency" in some pedagogic versions of classical theory.

[4] See J. Marschak and W. H. Andrews (1944), 145.

> *firm to firm and from year to year, depending on the technical knowledge, the will, effort, and lack of a given entrepreneur: these factors can be summarized as "technical efficiency," and may be represented by one or more random parameters.*

This is obviously true even judged by our present knowledge.

Technically, it was suggested that productive efficiency can be treated as a random parameter rather than a variable. In practice, however, the proposed random parameter model had never been as popular as some simplified models with non-random parameters. To capture the performance difference among the firms, a production function may have following form:[5]

$$u = A_r \, f(x_r) \qquad r = 1,2... \tag{3.1}$$

where **x** and **u** are input and output respectively and parameter A_r is regarded as a measure of technical efficiency for the r-th group of firms. Though the model does distinguish inter-group performance, it can not provide firm-specific efficiency measuring. This weakness is shared by all the models with a parameterized efficiency measure.

More conventionally, the effect of technical inefficiency may be captured by a stochastic variable

---

[5] The model by Lau and Yotopoulos (1971) should be regarded as a modern version of classic treatment.

randomizing the production function. Correspondingly, the production function can be written as[6]

$$u = f(x, \beta, \varepsilon) \tag{3.2}$$

where ε is a random term representing "technical inefficiency". The conventional econometric techniques are believed applicable to the model. Though the production function **f()** may have a complicated form to meet various well-behaved properties, the treatment of technical efficiency was usually quite simple in the early models. It was common practice to attach interpretation of inefficiency to the unexplainable estimation error. The efficiency of an observation was measured by its relative distance to the estimated function.

The most problematic part of the classical approach is its average treatment of the production function. In this approach, the potential output for a given input level is a statistical average one rather than a maximum. The estimated production function is, therefore, basically inconsistent with the theoretical definition. Furthermore, the economic significance of such measured technical efficiency is questionable. Though by comparing the relative position of an observation to the production function, one can rank a group

[6] See Marschak and Andrews (1944).

of observations ("firms"), it is frequently noted that the efficiency measure has only statistical significance rather than economic significance.[7]

These followed some new models which, though viewed as classical models, are pioneer works in frontier production function analysis. According to Klein (1953), if productive inefficiency is under consideration, the production function for profit maximizing firms is:

$$u = L^{\alpha} K^{\beta} \varepsilon$$
$$\alpha v_{\alpha} = \frac{z}{z_L}, \qquad \beta v_{\beta} = \frac{z}{z_K} \tag{3.3}$$

where **z** is the total revenue from sales of the product **u**; $\mathbf{z}_L$ and $\mathbf{z}_K$ are the firm's expenditures for the factors **L** and **K**; the **v** parameters represent failure of short-run profit maximization; and, finally, ε is a random error. Due to the presence of the **v** parameter, performances of firms vary according to their success in maximizing the profits. It should be noted that the **v** parameters are regarded as systematic components in production. One may further introduce random components into the model. The resulting

---

[7] As Schmidt (1977) stated, the parameter set obtained from OLS estimation of a production function are consistent except for the constant term, thus one can "rank-order firms by efficiency - the more positive or less negative the residual, the more efficient the firm ...".

model then has the form of

$$u = L^{\alpha \varepsilon_\alpha} K^{\beta \varepsilon_\beta} \varepsilon$$

$$\alpha\, v_\alpha = \frac{z}{z_L\, \varepsilon_\alpha}, \qquad \beta\, v_\beta = \frac{z}{z_K\, \varepsilon_\beta} \tag{3.4}$$

where ε's are firm-specific random components. As a result of introducing the random components, the production structure has a composite error kernel. According to Nerlove (1965), Klein's models in fact have a frontier property and efficiency measurement based on the models is exactly consistent with Farrell's efficiency structure. However, most of the classical models do not have this nice feature. In contrast, conventional classical treatment features an average production function which presents to be a major problem prohibiting them to have a correct measurement of the technical inefficiency.

3.2 Production Frontier Models

As demonstrated in the last Chapter, Farrell's measurement of technical efficiency

$$F(x^*, u^*) = \min \{ \tau \ ; \ \tau x^* \in L(u^*) \} \tag{3.5}$$

is one of the earliest production frontier models. In the model, the technical efficiency of an observation x^* is evaluated by its radial distance to a production frontier. This structure spawned a number of frontier production models in 1960's and 1970's.

Aigner and Chu (1968) proposed a programming procedure to obtain a parametric production frontier as:

$$\begin{gathered} \min_{\beta} : \ t \\ s.t : \ t = \ln f(x;\beta) - \ln u \\ t \geq 0 \end{gathered} \tag{3.6}$$

where **t** is the term of technical inefficiency and all other variables are defined in a conventional way. A programming procedure was suggested to solve for all the parameters and give a parametric description of the production function **f(x)**. Fundamentally, **f(x)** can keep a frontier property by imposing the restriction $\mathbf{u} \leq \mathbf{f(x)}$.

However, the parametric form can be a straitjacket if it is not flexible enough or mis-specified. In the model, observed input-output data are explained by two components: the part explained by the estimated parameter set and the residuals interpreted as inefficiency. Clearly, should the functional form of the production function be mis-specified, technical efficiency cannot be a precise measure.

Another problem associated with the above frontier model is its inability to handle random errors. Conceivably, the deviation from a frontier could be caused by either technical efficiency factor or other random factor. Should the random component be handled improperly, the production frontier could be mis-placed and efficiency measurement would be affected.[8] However, the pure random errors were often ignored in the early frontier models. On this negligence, Farrell acknowledged[9]:

> *... errors of observation will introduce an optimistic bias, which can only be eliminated if the distributions of both errors and efficiencies are known. .... for practical purposes the important fact is that if errors are small compared with the variation in efficiencies, this bias will be negligible.*

Therefore, the validity of the models depends on the

---

[8] An empirical example can be found in D. Deprins, L. Simar and H. Tulkens (1985). In their study of labour-efficiency in Belgian Regie des Postes, extraordinary outliers are found to have either "global" or "local" effects on the frontier and hence the estimated indexes.

[9]. M. J. Farrell (1957), 263.

presence of only small random factors.[10] In this regard, the econometric method should have some advantages.

Deterministic frontier models were the dominating statistical approaches in 1970's.[11] The typical deterministic frontier model has the following structure:

$$u = f(x;\beta)\cdot e^{-t} \qquad t \geq 0 \tag{3.7}$$

where variable **t** is technical inefficiency term and is subject to some pre-specified statistical distribution. It has been suggested that **t** may have a Beta [Afriat (1972)] or a Gamma [Richmond (1974), Greene (1980)] distribution. The suggested estimation techniques include corrected OLS method (COLS), which relocates the production function by correcting the constant term in order to ensure a frontier property, and the maximum likelihood method (ML), which requires a properly specified distribution for the efficiency term to yield consistent and asymptotically efficient estimation of the

---

[10] There are some other factors that may cause the observed efficiencies to differ from factual ones. Those factors have external effects on the production process. The weather condition is an example of such a factor. Those pure random factors may be assumed to have a normal distribution. See R. Forsund et. al.(1980)

[11] The approach was initially proposed by Afriat (1972).

parameter set.[12]

The mostly celebrated contribution of the deterministic frontier model, as it is called, is the frontier property. Now the production function keeps a perfect consistency with its theoretical definition, i.e., the maximum attainable output level for given inputs. However, the interpretation of technical inefficiency by the deterministic frontier approach is much less convincing. The question is: if the deviation from the frontier is viewed as technical inefficiency, there would be no room for the pure random noise. In this regard, strength of the econometric method is not fully employed in the deterministic frontier model. This deficiency turns out to be a major motivation for developing the more sophisticated stochastic frontier model.

---

[12] See Richmond (1974) for the COLS technique and Greene (1980) for the ML method and associated problems.

3.3 The DEA and the Stochastic Frontier Models

The following two models emerged in late 1970's as the two major approaches in production and efficiency analysis.

(1). The Stochastic Frontier Model.

Based upon the earlier deterministic frontier models, the stochastic frontier model features a composite error structure.[13] It has the following typical form:

$$u = f(x;\beta)\, e^{v-t}, \qquad -\infty < v < \infty, \qquad t \geq 0 \tag{3.8}$$

where **f( )** is a deterministic production function and **exp(v-t)** respresents the error structure. In the kernel of the error term, **v** is the random component and is conventionally assumed to have a normal distribution. Further, in the earlier versions of the stochastic frontier model, **t** is assumed to have either a half-normal or an exponential distribution. Moreover, it was assumed that **v** and **t** are statistically independent.

Let us define ε **= (v-t)**. Given the probability density functions (*pdf*) $p_v(v)$ and $p_t(t)$, the marginal *pdf* of the composite error ε can be obtained by integrating the joint

---

[13] The earliest composite error structure of productive efficiency can be traced back to Klein (1953).

density functions:

$$p_\varepsilon(\varepsilon) = \int_0^\infty p_t(t)\ p_v(\varepsilon + t)\ dt \tag{3.9}$$

This leads to the maximum likelihood function:

$$\begin{aligned} ML(\varepsilon) &= \prod p_\varepsilon(\varepsilon) \\ \textit{where} & \\ \varepsilon &= \ln u - \ln f(x;\beta) \end{aligned} \tag{3.10}$$

Jondrow et. al (JLMS, 1982) proposed a conditional estimation procedure to extract the technical efficiency term **t** out of the composite error ε. According to the Baysian rule, the *pdf* of **t** conditional on ε is given by:

$$p_t(t \mid \varepsilon) = \frac{p_{t,v}(t,\varepsilon)}{p_\varepsilon(\varepsilon)} \tag{3.11}$$

Finally, the expectation and mode of **t** conditional on ε can be obtained by:

$$\begin{aligned} E(t \mid \varepsilon) &= \int_0^\infty t\ p_t(t \mid \varepsilon)\ dt \\ M(t \mid \varepsilon) &= t \mid_{\frac{\partial p(t \mid \varepsilon)}{\partial t} = 0} \end{aligned} \tag{3.12}$$

The major attraction of the model comes from its subtle error structure. The other statistical models reviewed previously, though maintaining a frontier property, are not able to distinguish technical inefficiency and the effect of random noise. The frontier is therefore subject to the influence of stochastic disturbances. With the composite error structure, extraordinary behaviour of an observation can be filtered out and has less effect on the location of the production frontier.

The limitation, on the other hand, is the restrictive efficiency distribution form imposed as the model is specified. As many researchers realized, "the more structure we impose on a model the better our estimates - provided the structure we impose is correct."[14] However, correct specification can never be assured unless we have perfect knowledge of the real world. Thus specification error of the efficiency distribution presents a major threatening to the stochastic frontier approach. Furthermore, the correct specification of the functional form for the production process is of important. Should the functional form were mis-specified, the location and the shape of the production frontier would be interpretated improperly. Finally, the effects would be dumped to and distort the efficiency estimation.

(2) The Non-parametric DEA Model

Following Farrell's measurement scheme, Charnes,

---

[14] Bauer (1990), 40.

Cooper and Rhodes (1978) suggested that technical efficiency can be measured by solving a fractional programming problem

$$\max_{\alpha,\beta}: \quad \frac{\Sigma\alpha_j u_j^*}{\Sigma\beta_i x_i^*}$$

$$s.t.: \quad \frac{\Sigma\alpha_j u_j^r}{\Sigma\beta_i x_i^r} \le 1, \qquad i = 1, \ldots, n \qquad \textbf{(3.13)}$$

$$j = 1, \ldots, m$$

$$\alpha_j,\ \beta_i \ge 0 \qquad r = 1, \ldots, k$$

where **n, m** and **k** are the dimensions of input, output and the numbers of observations respectively.[15] Essentially, the programming procedure known as DEA (Data Envelopment Analysis) creates two convex cones in input and output space respectively. The efficiency of an observation wrapped in the cones can be measured by the radial distance from the data point to a matching point on the hull of the convex body. Clearly the procedure conforms to the Farrell's measurement of technical efficiency defined previously.

The distinguishing feature, the non-parametric property, of DEA can not be overstated. In DEA, one does not need to specify a functional form either for the production structure or for the efficiency structure. The risk of mis-specification in this regard is reduced correspondingly.

---

[15] There are a number of other variations of the basic DEA model which we will discuss in Chapter 4.

However, it is important to understand that the lack of a requirement on the functional form does not imply the non-existence of functional relationship in production. In fact, as we will discuss later, a DEA solution does defines a mapping from input to output space[16]. Further, the feasibility of the DEA analysis depends on only whether there are enough observations to span the convex cones in input and output space. This implies that the requirement for DEA in terms of on the sample size is much less restrictive than for statistical models.

The shortcoming of the DEA model is its inability to cope with random noise. As a non-statistical model, DEA considers all the deviations, including the ones caused by random noise, from the data envelope as systematic inefficiency.[17] The consequence is: extraordinary behaviour of any observation may exert a direct impact on the location of the production frontier and hence on efficiency measurement. Should the random noise present in production, the DEA tends to under-estimate technical efficiency.

Another impending risk to the DEA model is mis-specification on the returns to scale property. Though a functional form specification is not required, as we will

---

[16] See the discussion in Chapter 4.

[17] In this regard, Farrell's statement quoted in the last section remains valid.

demonstrate in Chapter 4, specification of a returns to scale property is a compulsory requirement for the DEA. The returns to scale property defines the shape of the production frontier, and hence determines the interpretation of entire data set and the measurement of the technical efficiency. Therefore, accuracy of efficiency measurement cannot be ensured unless the returns to scale property is correctly specified.

3.4 Recent Development of Production Frontier Analysis

The recent period is one of elaboration, in which both the stochastic frontier model and the DEA model continue their domination in efficiency analysis. Some important developments are worthy noticing in both approaches.

On the stochastic frontier analysis, the major efforts have been directed to three major dimensions.[18]

First, much effort has been paid to seeking more flexible functional forms for the composite error production structure. Research along this dimension follows a similar pattern to that we have seen in traditional production

---

[18] For a detailed review of recent developments in the stochastic frontier approach, see Bauer(1990).

analysis. The cost function has been a primary vehicle used to reveal the production and efficiency structure. In addition, the family of share equations is brought into the analysis and has been modelled with a composite error structure. The stochastic frontier model has been extended to cover the entire cost system. In this cost system approach, one may find more detailed discussions in Greene (1980) and Kumbhakar (1989).

The second dimension is to relax the less flexible structure imposed on the (either technical or allocative) efficiency terms. The dominating models of error structure since late 1980's have been the half-normal (technical efficiency)/normal (random noise) and exponential (technical efficiency)/normal (random noise) models proposed by ALS (1977). Based on these two models, JLMS (1982) developed the conditional estimation technique. As a result, the two models dominated the stochastic frontier approach for the rest of 1980's. However, there is no emperical evidence suggesting that the efficiency distribution follows those particular distribution forms. Stevenson (1980) developed two generalized models: the truncated-normal (technical efficiency)/normal (random noise) model and Gamma/normal model. However, the conditional estimation procedures for these two generalized models were not worked out until recently due to the complexity of the mathematical structure.

Recently, Greene (1990) has developed the conditional estimation procedure for the Gamma/normal model. With a flexible Gamma distribution, greater freedom can be achieved for the efficiency structure.

The third dimension of advance in the stochastic frontier approach is the application to the panel data. The primary advantage of using panel data is to obtain consistency of efficiency estimation which is not available in cross section analysis. Moreover, using panel data may reduce the reliance on the restrictive assumption on the efficiency distribution. Finally, an industry's efficiency may have a time profile in addition to a cross section profile [see Jondrow et. al. (1982), Schmidt and Sickles (1984), Kumbhakar and Summa (1989)].

New developments in estimation techniques have also appeared recently. While maximum likelihood estimation serves as the major vehicle, alternative techniques are also applied. Noticeably, the moment method is more frequently employed in estimation. To estimate a stochastic frontier function

$$y = f(x;\beta)\, e^{v-t} \qquad \textbf{(3.14)}$$

the core issue is to find out the parameter sets for the probability density functions of **v** and **t**. In this regard, the various order moments of the composite error, i.e., $\mathbf{E[(v-t)^r]}$, impose many restrictions (therefore provide a great amount of information) on the parameters. From these moment restrictions, the parameter set can then be identified. Because the composite error usually has a cumbersome complex functional structure, the moment method has an obvious advantage in revealing the structure. Regarding empirical application, Greene (1990) suggests using the moment method to estimate the proposed Gamma/normal error structure. Kopp and Mullahy (1990) discusses various aspects of applying the moment method to the stochastic frontier model. They suggest that the moment restrictions may play an important role in testing the stochastic frontier model specification.

On the side of the non-parametric approach, we also find some important advancement. The most important achievement in the DEA approach is the polyhedral cone-ratio generalization of the DEA models. In the basic DEA model, the referenced efficiency set is obtained inclusively from the data set under investigation. This referenced input-output sets define a convex cone in both input and output spaces. Efficiency can then be measured as an input-output cone ratio. However, according to Charnes, Cooper, Wei, and Huang (1986) and Charnes, Cooper, Huang and Sun (1990), the cones can be defined much more generally so that external information can be employed in the cone construction. As a result, the external evaluation for the input and output could be imposed as a priori weight (or multiplier) vector to assist the cone building. This new modelling technique has been applied into an empirical investigation by Thompson, Langemeier, Lee, Lee and Thrall (1990).

Another development in the DEA approach is the stochastic DEA model proposed by Sengupta (1988) (1990). In the standard DEA approach, individual efficiency estimation seems to be the ultimate intention. However, according to Sengupta, a statistical distribution can be imposed to fit the DEA efficiency estimates and provides an interface for the DEA and the statistical model.

The recent developments signify somewhat important changes in the methodology of production theory. These changes in turn influence the further development of efficiency study.

The most important, the change in methodology should be observed. Optimizing behaviour is a pivotal assumption in economic analysis. The convention of economics is to analyse the optimized track of economic activity, with the assumption that the track can utltimately be attained. Non-optimal or sub-optimal behaviour, except it as results from imperfect markets, is out of consideration. Following this convention, economic agent is assumed born efficient in production. Correspondingly, the production function estimated by the conventional econometric model describes "average" optimizing behaviour. Deviations from the function have only statistical significance rather than economic significance [Varian (1990)]. The surging interest in frontier production models indicates that the economic significance of violating an optimizing model is now becoming a major concern of production analysis.

It is worth noting too that trend of efficiency study has been influenced by the changing methodology. The first noticeable change in the efficiency study is the firm-specific oriented analysis. Heterogeneous performances of firms had been observed by empirical production study for many years [e.g., Johanson (1970)]. However, econometrics had done little to model the phenomenon. Even for efficiency study, the focus was on the average efficiency of an industry. The stochastic frontier model, and especially the conditional estimation technique has changed the situation. Currently, the firm-specific efficiency analysis becomes a new standard for efficiency study. The theory underlying for this orientation is that the economic agents (or DMU -decision making unit) are heterogeneous. The role of efficiency study is to distinguish rather than blur the differences. This thinking brings out a convergency of interests of microeconomics and management science regarding efficiency analysis.

Further, the study of the industrial efficiency profile characterizes another aspect of recent developments. The industrial efficiency profile is often referred to as the efficiency distribution of an industry. There are many reasons to direct our attention to this rarely explored aspect. Empirical studies suggest that, even using same input factors and producing same outputs, firms may have diversified performance.[19] Many factors account for the performance discrepancy: technology invention and diffusion, geographic distribution, equipment replacement and upgrading. As a result, each industry features its own pattern of efficiency profiles.

The earliest effort to model the efficiency profile can be dated to Farrell (1957) who gave a first description of efficiency distribution. In a more formal way, Afriat (1972) proposed a Beta distribution to characterize the efficiency profile. In fact, all efforts in stochastic frontier approaches are focused on one point: seeking a flexible distribution form to capture the underlying efficiency

---

[19] As Johansen (1972) observed, the performance of some inputs may have a difference factor of two to four times. A more recent example can be found in Deprins, Simar and Tulkens (1984). By analyzing the labour efficiency structure in the Belgian Post Office, they found that there is a wide spread spectrum of possible efficiency profiles.

profile.[20] We observed that recently the ability of recovering the efficiency profile has became a criterion to assess alternative estimation procedure [See Bjurek and Hjalmarsson (1990)].[21]

3.6 Summary

In this chapter, we reviewed alternative modelling techniques for efficiency estimation. From this review, we found that production frontier analysis is the central issue of efficinecy measurement. In fact, two production frontier models, i.e., the stochastic frontier model and the DEA model are leading the current efficiency studies. In addition, we also discussed recent developments and their siginificances. With these understandings, we will devote our efforts to these two leading approaches in the next two chapters.

---

[20] The concept of "capacity distribution" propsoed by Johansen (1972) bears a similar meaning. For recent concern about the concept, see Muysken (1985).

[21] As matter of fact, to handle an irregular pattern of efficiency distribution (e.g., a skewed one) is one of major motivation to apply the SF model to panel data. It is hoped that reliance on the restrictive distribution form of efficiency can be reduced by introduing time dimension. See Kumbhakar and Summa (1989).

CHAPTER 4. NON-PARAMETRIC ESTIMATION OF TECHNICAL EFFICIENCY: A PROGRAMMING APPROACH.

This chapter concentrates on one of the leading approaches to efficiency estimation: the non-parametric programming approach. The approach introduced by Farrell aimed to solve the efficiency estimation problem in a CRS (constant return to scale) production process. Two important advances emerged in the 1970's which generalized Farrell's model. The analysis by S.N. Afriat (1972) focuses on theoretical aspects of efficiency estimation. In his analysis, a series of "inner envelope" propositions can be regarded as the theoretical foundation for the programming approach. Further, it serves as a bridge linking this approach with other conventional approaches. Later, in the management science literature, R. Banker, R. Charnes and W. Cooper (1978) formulated the DEA problem, as we outlined previously, which greatly enhanced the non-parametric programming approach.

This Chapter attempts to re-interpret the non-parametric approach in a way more consistent with economic analysis. The chapter starts with a discussion of the properties associated with a data dependent production function (DDPF). This discussion is applicable to empirical production frontier analysis in general. In section two, we discuss a set of linear production frontiers and, following Afriat (1972), and provide a second proof for a series of "inner envelope" propositions pertaining to these frontiers. We demonstrate that these propositions play a pivotal role in linking parametric and non-parametric production frontier models. Based on these propositions, section three reviews two sets of efficiency estimation models. These two sets of models are in fact the primal and dual formulations of one linear programming (LP) problem. Then, in section four we will trace out the technology correspondence in the duality relationship. The returns to scale property is a very important consideration should a non-parametric efficiency estimation model be employed. The issue is discussed in section four together with some other issues. The chapter ends with a summary.

4.1 The Data Dependent Production Frontier

This section addresses the notion of the efficient production frontier and its interface with empirical data.

Assume there are **k** observations in a set of input-output data. The activities of the data set (or, activity of an industry) can be characterized by an input matrix

$$X = \begin{bmatrix} x_1^1 & x_1^2 \ldots & x_1^k \\ \ldots & \ldots & \ldots \\ x_n^1 & x_n^2 \ldots & x_n^k \end{bmatrix}$$

and an output matrix:

$$U = \begin{bmatrix} u_1^1 & u_1^2 \ldots & u_1^k \\ \ldots & \ldots & \ldots \\ u_m^1 & u_m^2 \ldots & u_m^k \end{bmatrix}$$

A data dependent production frontier (**DDPF**) can be jointly defined by these **k** observations.

More specifically, a **DDPF** describes a point to point mapping from input to output space:

$$f : x \rightarrow u$$

In this relationship we assume that the previously outlined axioms about **L(u)** hold, and **f** is the boundary of the graph **GrL** describing the data set. It represents the frontier of a technology transforming input into output.

As a function constructed from the observed data set, a **DDPF** features the following characteristics:

(1) *Theoretical Consistency.*

The **DDPF** should interpret the observed data consistently with all the axioms **L.1** through **L.7** introduced in Chapter 2. and in particular:

(a) To be consistent with **L.4** and **L.5**, **f** is a non-decreasing function on the domain of empirical data **D(X)**,

$$\mathbf{D(X)} \equiv \{ \mathbf{x}: \ \min[x_i^r] \leq x_i \leq \max[x_i^r], \ \mathbf{x} \in R_+^n, \quad r=1,\ldots,k \}.$$

and on an extended domain of the empirical data:[1]

$$\mathbf{ED(X)} \equiv \{ \mathbf{x}: x_i \leq \max[x_i^r], \ \mathbf{x} \in R_+^n, \ r=1,\ldots,k \}.$$

(b) By **L.7**, **f** is concave either on **ED(X)** or on D(X).[2]

(c) Since **f** is the boundary of graph **L**, it is at least weakly efficient.

(2) *Data Consistency*

A **DDPF f** is consistent with a given set of data in the sense that at least for some observations **(x',u')**, we have

$$\mathbf{u'} = \mathbf{f(x')}$$

and for all other observations **(x,u)**,

$$\mathbf{u} < \mathbf{f(x)}$$

It is crucial important for a production function to have the frontier property. Should this property and all the axioms be satisfied, the production function **f** is an efficient frontier.

Data consistency makes technical inefficiency a

---

[1] The extended domain includes the origin.

[2] However, strict concavity is not required.

legitimate part of the production structure. Deviation from the production frontier is regarded as a permissible behaviour. In the non-parametric approach, the performance gap

$$\mathbf{f(x) - u}$$

bears the systematic interpretation of inefficiency. However, in the statistical production models, the gap may be assumed to contain both a systematic and a purely random component.[3]

(3) *Requirement for Sufficient Observations*

To ensure the existence of a DDPF in space $R_+^m x R_+^n$, it is necessary to require that the number of observations should not be less than the sum of the dimension of inputs and outputs.[4] If this requirement were not satisfied, the production possibility set would degenerate (at least in some dimensions), and we would not be able to estimate DDPF **f**. However, this requirement is a necessary rather than a sufficient condition for the existence of an empirical

---

[3] According to Farrell (1957), the stochastic component, if it does exist, can be reasonably ignored provided it is relatively small.

[4] Strictly speaking, the requirement pertains to the construction of the production frontier in space R_+^{m+n}. Seiford and Thrall (1990) states that this is also the condition for the DEA model. As matter of fact, the condition for the DEA model should be $k \geq max[m,n]$. This is the necessary condition to span production frontiers in both input and output space respectively.

frontier. In the programming approach, linear dependence between different observations may still cause the degeneration of a frontier.

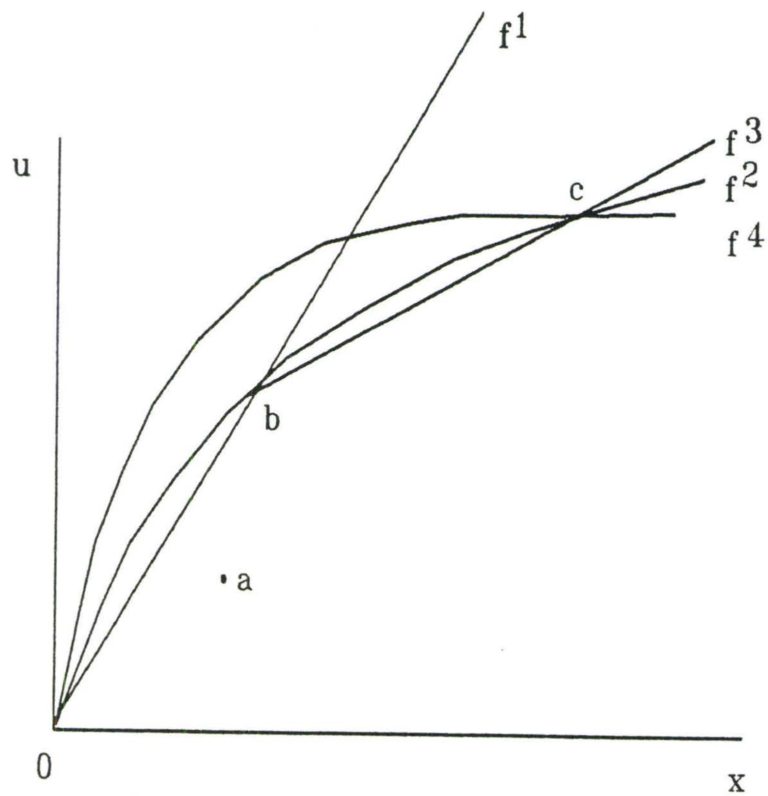

Figure 4.1 A family of DDFP

The above characteristics of DDPF are illustrated in

Figure 4.1. For a given set of observations such as **(a, b, c)**, a family of production frontiers satisfying the above characteristics can be constructed. The form of these candidates may vary from linear, e.g. f^1 and f^3 (o-b-c), to non-linear, e.g. f^2 (o-b-c) and f^4 (o-c). The frontier f^1 is ray (CRS) production function which passes through the data point **b** and leaves observations **a** and **c** as interior points of the frontier. The frontier f^3 is a linear function with a non-increasing return to scale **(NIRS)** property. It has a **CRS** portion **(o-b)** and a portion **(b-c)** with a decreased marginal product. The **NIRS** frontier goes through both observation **b** and **c**, and thus both are identified as efficient observations. f^2 also goes through the observation **b** and **c**, but the frontier is non-linear. f^4 is another possible non-linear frontier to interpret the given data set. However, only observation **c** is efficient if f^4 is referenced. All of these frontiers are in fact valid DDPFs and may be employed to interpret the data set. It is easy to check that all of these frontiers maintain theoretical consistency with our previously outlined axioms. Also, all of them are featured with data consistency by covering the observations in one way or other. This example shows that the DDPF for a given data set usually refers to a family of possible candidates.

The propositions in the next section prove that, in each returns to scale category, the family of frontier production functions is bounded from below by a piecewise linear production frontier. Those linear frontiers are in fact a set of "inner envelopes" or "underwear" covering the body of a given data set.

4.2 The Inner Envelope Propositions

In this section, we introduce three data dependent linear production frontiers. The efficient "inner envelope" property is stated by a series of propositions, which are based closely upon the theorems proposed by Afriat (1972).

Denote $\lambda = (\lambda^1,...,\lambda^k)$ as an intensity vector evaluating the observations' participation in frontier building. Three types of frontiers can be defined by imposing further restrictions on the intensity vector[5] : Constant Return to Scale (CRS), Non-Increasing Return to Scale (NIRS) and Variable Return to Scale (VRS).

4.2.1 CRS Production Frontier

Of many production frontiers satisfying the declared properties, a piecewise linear model featuring CRS technology is expressed as

---

[5] The technologies are classified according to the their global property, which refers to a model's behaviour in the domain **ED(X)**.

$$f^{c}(x) = \max_{\lambda} \left\{ U\lambda \, : \, x \geq X\lambda, \; \lambda \in R_{+}^{k} \right\} \tag{4.1}$$

Moreover, it has an inverse which can be expressed as:

$$g^{c}(u) = \min_{\lambda} \left\{ X\lambda \, : \, u \leq U\lambda, \; \lambda \in R_{+}^{k} \right\} \tag{4.1.a}$$

(1) *"Inner Envelope" Property*

The following proposition (due to Afriat) states that (4.1) represents a technically efficient CRS technology and has an "inner envelope" property.

Proposition 4.2.1 *If there exist k activities with inputs* $\mathbf{X} \in \mathrm{R}^{n}_{+}$ *, outputs* $\mathbf{U} \in \mathrm{R}^{m}_{+}$ *and axioms* **L.1-L.7** *being satisfied, then the linear mapping* $\mathrm{f}^{c}:\mathrm{R}^{n}_{+} \rightarrow \mathrm{R}^{m}_{+}$ *represents an efficient CRS technology (homogeneous degree 1) and serves as a lower bound of any other efficient CRS mapping* $\mathrm{f}:\mathrm{R}^{n}_{+} \rightarrow \mathrm{R}^{m}_{+}$.

Proof: Assume there is some other data dependent **CRS** frontier, say **f()**, satisfying the claimed axioms[6], then the non-decreasing property of the production frontier requires:

$$f(x) \geq f\left(\sum \lambda^{r} x^{r} \right), \; if \; x \geq \sum \lambda^{r} x^{r}, \quad \lambda^{r} \geq 0, \; r = 1, \ldots, k \tag{4.2}$$

---

[6] A single-output/multi-input production process is assumed in the proof for the sake of convenience, though the proof can be extended to cover multi-input, multi-output cases without much difficulty.

The homogeneity property of the CRS production function and the convexity axiom stated by **L.6** imply that **f()** is a concave function featured by[7]

$$f(\Sigma\lambda^r x^r) \geq \Sigma\lambda^r f(x^r), \quad \lambda^r \geq 0,\ r = 1,\ldots,n \tag{4.3}$$

Further, since **f** is non-decreasing, we have

$$f(x) \geq \Sigma\, \lambda^r f(x^r), \ if\ x \geq \Sigma\lambda^r x^r, \quad \lambda^r \geq 0,\ r = 1,\ldots,k \tag{4.4}$$

This relation can be rewritten as:

$$f(x) \geq \{\, \Sigma\, \lambda^r f(x^r), \ x \geq \Sigma\, \lambda^r x^r, \quad \lambda^r \geq 0,\ r = 1,\ldots,k\,\} \tag{4.5}$$

and further, in particular, we have[8]:

---

[7] For a concave function **f**, we have

$$f(\Sigma\frac{\lambda^r}{\Sigma\lambda^r}\, x^r) \geq \Sigma\frac{\lambda^r}{\Sigma\lambda^r} f(x^r) \qquad \lambda^r \geq 0$$

Further, if **f** is homogeneous of degree 1, we are able to get (4.3) by cancelling out the common denominator. Thus, CRS production function **f** has a cone property [see Nikaido (1972), 188].

[8] Since (4.5) holds for any λ, it holds for the particular λ that satisfying (4.6).

$$f(x) \geq \max_{\lambda}\{ \sum \lambda^r f(x^r), \; x \geq \sum \lambda^r x^r \quad \lambda^r \geq 0, \; r = 1,...,k \} \tag{4.6}$$

Consider the right hand side of **(4.6)**. Since **f(x)** is data consistent, for some observations **(x',u')**, we have

$$\mathbf{u' = f(x')}$$

and for all other **(x,u)**

$$\mathbf{u < f(x)}$$

Thus for non-negative λ, we have

$$\sum \lambda^r f(x^r) \geq \sum \lambda^r u^r \tag{4.7}$$

Therefore, we have:

$$\begin{aligned} f(x) &\geq \max_{\lambda}\{ \sum \lambda^r f(x^r), \; x \geq \sum \lambda^r x^r \\ &\qquad \lambda^r \geq 0, \; r = 1,...,k \} \\ &\geq f^c(x) = \max_{\lambda}\{\sum \lambda^r u^r, x \geq \sum \lambda^r x^r, \\ &\qquad \lambda^r \geq 0, r = 1,...,k\} \end{aligned} \tag{4.8}$$

This proves the "inner envelope" and efficient properties of $\mathbf{f^c(x)}$.

Finally, it can be verified that $\mathbf{f^c(x)}$ has the CRS (homogeneous of degree 1) property as follows. Let δ be a positive number, then[9]

$$
\begin{aligned}
f^{c}(\delta x) &= \max\left\{ U\lambda : \delta x \geq X\lambda,\ \lambda \in R_{+}^{k} \right\} \\
&= \max\left\{ \delta\, U\lambda : \delta x \geq \delta X\lambda,\ \delta\lambda \in R_{+}^{k} \right\} \\
&= \max\left\{ \delta\, U\lambda : x \geq X\lambda,\ \lambda \in R_{+}^{k} \right\} \\
&= \delta f^{c}(x) \quad .
\end{aligned} \tag{4.9}
$$

Thus the homogeneity property is proved.

Q.E.D.

---

[9] The following properties of $\mathbf{f^{c}(x)}$ should be noted. Let α be a positive real number, then

$$
\begin{aligned}
f^{c}(x) &= \max_{\lambda}\left\{ U\lambda\alpha : x \geq X\lambda\alpha, \lambda\alpha \in R_{+}^{k} \right\} \\
&= \max_{\lambda}\left\{ U\lambda : x \geq X\lambda,\ \lambda \in R_{+}^{k} \right\}
\end{aligned}
$$

i.e. scale of intensity vector does not affect the function evaluation. However, we should note

$$
\begin{aligned}
&\max_{\lambda}\left\{ U\lambda\alpha : x \geq X\lambda, \lambda \in R_{+}^{k} \right\} \\
&= \alpha \max_{\lambda}\left\{ U\lambda : x \geq X\lambda,\ \lambda \in R_{+}^{k} \right\}
\end{aligned}
$$

in the linear relationship. These two points are fairly transparent if $\mathbf{f^{c}(x)}$ is rewritten in linear programming problem form as in (4.10).

(2) *The nature of the solution*

The linear frontier $\mathbf{f}^c(\mathbf{x})$ is in fact the solution to an optimization problem with the intensity vector λ as the choice variables. The process selects some observations such as $(\mathbf{x}',\mathbf{u}')$ to build a frontier by assigning them a positive λ. For all other interior points such as $(\mathbf{x},\mathbf{u})$, the intensity elements would be assigned a value of $\lambda = 0$.

To demonstrate the nature of the optimal solution that characterizes (4.1), we rewrite it in the form of a linear programming (LP) problem:

$$\begin{aligned} &\max_{\lambda} : \sum_r \lambda^r u^r \\ &s.t. : \sum_r \lambda^r x_i^r \le x_i, \quad i = 1,\dots,n \\ &\lambda^r \ge 0 \end{aligned} \tag{4.10}$$

This can be further expressed by the Lagrangian function:

$$\max_{\lambda} \min_{\delta} L(\lambda,\delta) = \sum_r \lambda^r u^r + \sum_i \delta_i (x_i - \sum_r \lambda^r x_i^r) \tag{4.11}$$

where δ's are Lagrangian multipliers (and the shadow prices for the primal). The above Lagrangian function can be rewritten as:

$$\max_{\lambda} \min_{\delta} L(\lambda,\delta) = \sum_i \delta_i x + \sum_r \lambda^r (u^r - \sum_i \delta_i x_i^r) \tag{4.12}$$

This, in fact, corresponds to the dual problem:

$$\begin{aligned} &\min_{\delta}: \sum_i \delta_i x_i \\ &s.t.\ : \sum_i \delta_i x_i^r \geq u^r \\ &\delta^r \geq 0 \end{aligned} \tag{4.13}$$

The Goldman and Tucker theorem [Nikaido (1970), 252-254] states that for this type of programming problem, the optimal solution to the dual is characterized by the following conditions:

$$\begin{aligned} &(1) \quad u^r - \sum_i \delta^*_i x_i^r \leq 0\ , && r = 1, \ldots, k \\ &(2) \quad \sum_r \lambda^{r*}\left(u^r - \sum_i \delta^*_i x_i^r \right) = 0, && i = 1, \ldots, n \end{aligned} \tag{4.14}$$

where λ^* and δ^* are the optima. The first condition above implies that we can define

$$f(x) \equiv \sum_i \delta^*_i x_i \geq u \tag{4.15}$$

and the second implies:

$$\begin{aligned} &f(x^r) > u^r \quad \Rightarrow \quad \lambda^{r*} = 0, \\ &\lambda^{r*} > 0 \quad \Rightarrow \quad f(x^r) = u^r. \end{aligned} \tag{4.16}$$

Therefore, the optimization process ensures data consistency and that only some observations such as **(x',u')** are selected to build a **DDPF** of the form:

$$f^c(x) = \max_{\lambda} \{ \Sigma \lambda^r u^r, \; x \geq \Sigma \lambda^r x^r, \quad \lambda^r \geq 0, \; r = 1, \ldots, k \} \tag{4.17}$$

This again verifies the efficiency property of $\mathbf{f}^c(\mathbf{x})$.

This demonstration of the optimal status should be considered as an extension to the proof of the Proposition 4.1. It is extremely important to note that we had in fact demonstrated the relationship between the non-parametric frontier and a parametric frontier. Though the parameteric frontier associated with (4.15) is only a facet of a complete frontier, it bears a conventional interpretation of production function.

The linear production frontier $\mathbf{f}^c(\mathbf{x})$ is illustrated in Figure 4.2 where a facet denoted as f^c is spanned by the observations **(x',u')**. In fact a set of similar facets can be generated to wrap the entire data set. Obviously, any other production frontier like the nonlinear shell in the picture is bounded from below by this set of facets.

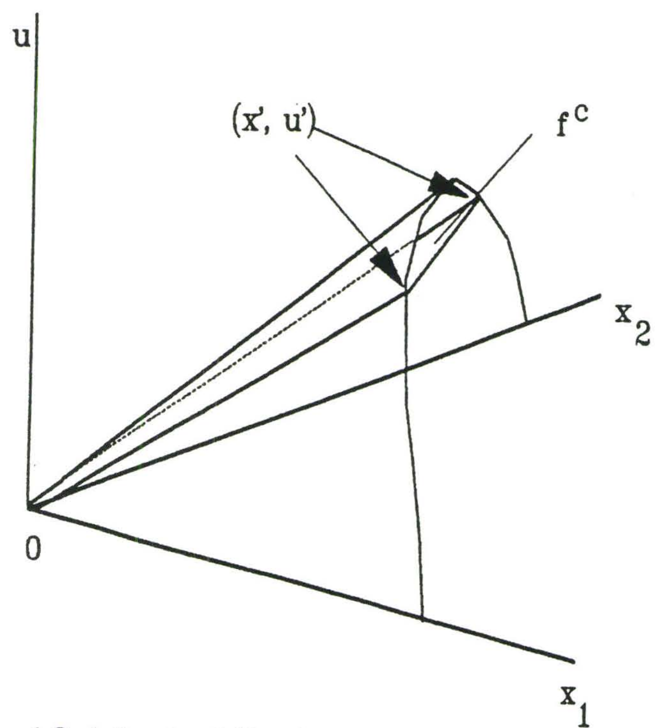

Figure 4.2 A facet of the inner envelope of the CRS frontier

4.2.2 NIRS Production Frontier

To illustrate the case of NIRS, let us consider data set **(a,b,c,d)** in Figure 4.3 (disregard **c'** and **d'** for the moment). As explained previously, we can construct a CRS production frontier along the ray **0-a-a'**. However, a production process may not keep the proportional expansion path after some production scale is reached. The marginal product may decline. To accommodate this situation, a linear production frontier may be constructed as **0-a-b** in Figure 4.3. This is a piece-wise linear NIRS production frontier.

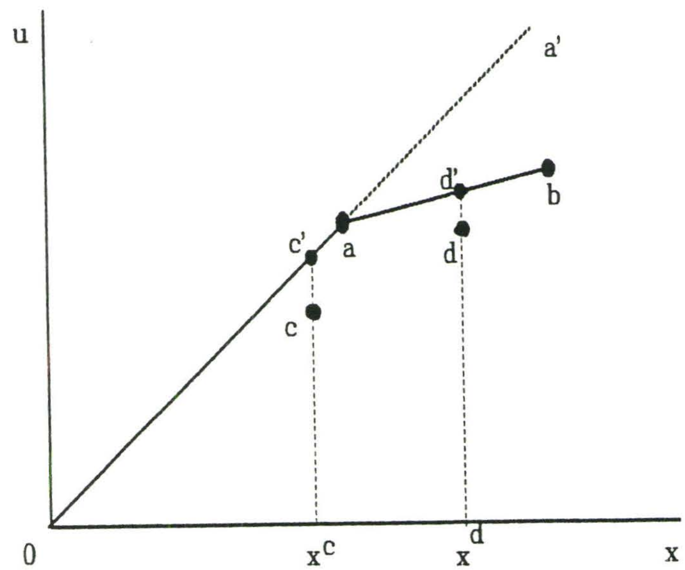

Figure 4.3 Linear NIRS production frontier

Due to the concavity assumption, the observations **c** and **d** can not be on the frontier. The potential outputs for the input level $\mathbf{x}^c$ and $\mathbf{x}^d$ are indicated by **c′** and **d′** respectively. To express these potentials by the given observations, we have:

$$c' = \lambda^0\, 0 + \lambda^a\, a, \qquad \lambda^0 + \lambda^a = 1$$
$$d' = \lambda^a\, a + \lambda^b\, b, \qquad \lambda^a + \lambda^b = 1$$

where λ's are the intensity (or weight) factors. It should be stressed that though the origin **f(0) = 0** now serves as a data point spanning the frontier, its appearance is usually nullified in conventional expressions. As a result, the restriction on the intensity vector becomes:

$$\Sigma\lambda^r \leq 1, \quad \lambda^r \geq 0, \quad r = 1, \ldots, k$$

This result leads to a linear model featuring the NIRS property:

$$f^n(x) = \max_{\lambda} \left\{ \lambda U : x \geq \lambda X, \ \Sigma \lambda^r \leq 1, \ \lambda \in R_+^k \right\} \qquad \textbf{(4.18)}$$

and it has an inverse:

$$g^n(u) = \min_{\lambda} \left\{ \lambda X : u \leq \lambda U, \ \Sigma\lambda^r \leq 1, \lambda \in R_+^k \right\} \qquad \textbf{(4.18.a)}$$

To show the efficiency and the "inner envelope" properties, we have

Proposition 4.2.2 *If there exist k activities with input* $\mathbf{X} \in \mathbf{R}^n{}_+$ *and* $\mathbf{U} \in \mathbf{R}^m{}_+$ *and the axioms* **L.1** *to* **L.7** *are satisfied, then a linear mapping* $f^n: R^n{}_+ \rightarrow R^m{}_+$ *expressed as (4.18) represents an efficient NIRS technology and serves as a lower bound of any other efficient NIRS mapping* $f^n: R^n{}_+ \rightarrow R^m{}_+$.

Proof. Assume there is some other production frontier **f()** satisfying the axioms **L.1 - L.7.** Further, **f** is concave on **ED(X)**, then

$$\sum \lambda^r f(x^r) \leq f(\sum \lambda^r x^r), \ \sum \lambda^r = 1, \quad \lambda^r \geq 0, \ r = 0,1,\ldots,k \tag{4.19}$$

Note that superscript **r** running from **0** to **k** so that **f(0) = 0** is included in the data domain. To express it explicitly, we have

$$\sum \lambda^r f(x^r) + \lambda^o f(0) \leq f(\sum \lambda^r x^r + \lambda^o 0), \quad \sum \lambda^r + \lambda^o = 1, \ \lambda^r \geq 0, \ r = 1,\ldots,k \tag{4.20}$$

Since **f(0) = 0**, we have

$$\sum \lambda^r f(x^r) \leq f(\sum \lambda^r x^r), \ \sum \lambda^r \leq 1, \quad \lambda^r \geq 0, \ r = 1,\ldots,k \tag{4.21}$$

Thus **(4.21)** is the condition that the linear piecewise **f** is concave on **ED(X)**. For the restriction on the intensity vector, equality sign must hold for all sections of the frontier except the portion adjacent to the origin. For the portion of the frontier adjacent to **f(0) = 0**, the summation of intensity factors λ is less than one since the **f(0) = 0** is only an implicit data point and its weight λ^o is not counted among the elements of intensity vector which has **k** elements.

Further, **f** is non-decreasing by axiom **L.4**, i.e.

$$f(x) \geq f(\Sigma\lambda^r x^r),\ if\ x \geq \Sigma\lambda^r x^r, \quad \lambda^r \geq 0,\ r = 1,...,k \tag{4.22}$$

If combining this with the concavity condition, we have

$$f(x) \geq \Sigma\lambda^r f(x^r),\ if\ x \geq \Sigma\lambda^r x^r, \quad \Sigma\lambda^r \leq 1\ ,\ \ \lambda^r \geq 0,\ r = 1,...,k \tag{4.23}$$

This can be rewritten as:

$$f(x) \geq \{\ \Sigma\lambda^r f(x^r),\ x \geq \Sigma\lambda^r x^r, \quad \Sigma\lambda^r \leq 1,\ r = 1,...,k\}. \tag{4.24}$$

in general and, in particular, we have

$$f(x) \geq \max_{\lambda}\{\ \Sigma\lambda^r f(x^r),\ if\ x \geq \Sigma\lambda^r,\ \lambda^r \geq 0 \quad \Sigma\lambda^r \leq 1,\ r = 1,...,k\} \tag{4.25}$$

Finally, as shown in the proof of the proposition **4.2.1**, if **f()** is data consistent, we have:

$$f(x) \geq \max_{\lambda} \{ \Sigma\lambda^r f(x^r), \; if \; x \geq \Sigma\lambda^r x^r, \; \lambda^r \geq 0$$
$$\Sigma\lambda^r \leq 1, \; r = 1, \ldots, k \} \qquad \textbf{(4.26)}$$
$$\max_{\lambda} \{ \Sigma\lambda^r u^r, \; if \; x \geq \Sigma\lambda^r x^r,$$
$$\lambda^r \geq 0, \Sigma\lambda^r \leq 1, r = 1, \ldots, k\}$$

This proves the efficiency and "inner envelope" property of $\mathbf{f}^n(\mathbf{x})$.

Q.E.D

The characteristics featuring the optimal solution of **(4.18)** can be found in a way parallel to what we have done for the **CRS** Case.

4.2.3 VRS Production Frontier

The third model has a variable returns to scale property. The VRS assumption allows the marginal product to increase in the first phase of production and then decrease after some production scale is reached. The Figure 4.4 illustrates the VRS case in a linear process. Again, we have a set of observations **a,b,c** and **d** as in the NIRS case. However, observation **c** now may serve as a point on the frontier under the VRS assumption. Should this be the case, the linear VRS frontier is **c-a-b**. It should be noted that if observation **c** (which is the closest to the origin among the all observations) is allowed to participate the frontier building, the origin could not be part of the frontier. Otherwise, the concavity condition stated in Chapter 2 for the production frontier would be breached. Therefore, a VRS linear frontier virtually does not cover the section between the origin and its most adjacent observation. In Figure 4.4, the VRS production frontier is **c-a-b** accordingly. Formally, the linear VRS model is concave just on **D(X)**.[10] The key point to formulate the VRS model is, therefore, to enforce the absence of the origin in the frontier building. The enforcement is handled by imposing a tighter constraint on the intensity vector, namely

$$\Sigma\lambda^{r} = 1, \quad \lambda^{r} \geq 0, \quad r = 1,\ldots,k \qquad \textbf{(4.27)}$$

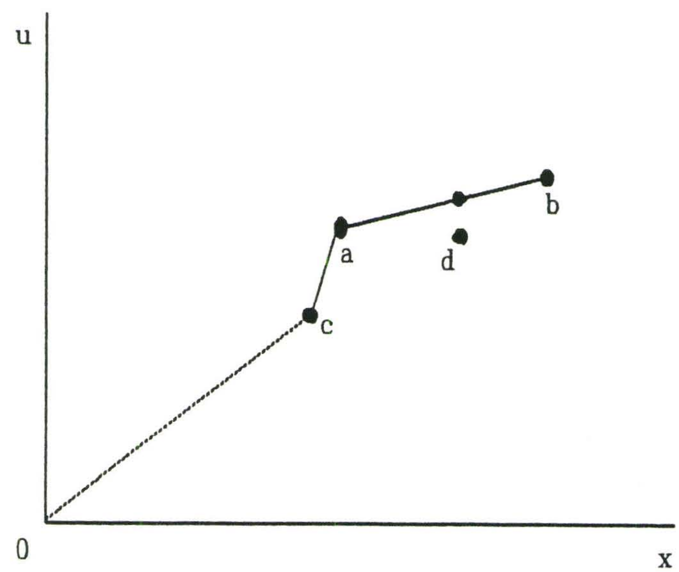

Figure 4.4 Linear VRS production frontier

[10] This assumption, however, is not consistent with the axiom **L.3** as there, the origin **f(0) = 0** has to be excluded from the graph **L(u)**.

on the intensity vector. The constraint brings out a linear VRS model with a form of

$$f^{v}(x) = \max_{\lambda} \left\{ \lambda U : x \geq \lambda X, \ \Sigma \lambda^{r} = 1, \ \lambda \in R_{+}^{k} \right\} \quad \textbf{(4.28)}$$

and it has an inverse:

$$g^{v}(u) = \min_{\lambda} \left\{ \lambda X : u \leq \lambda U, \ \Sigma \lambda^{r} = 1, \lambda \in R_{+}^{k} \right\} \quad \textbf{(4.28.a)}$$

To show the "inner envelope" property, we have

Proposition 4.2.3 *If there exist k activities with input* $\mathbf{X} \in \mathrm{R}^{n}_{+}$ *and output* $\mathbf{U} \in \mathrm{R}^{m}_{+}$ *and the axiom* **L.1 - L.2, L.4 - L.7** *are satisfied, then the linear mapping* $f^{v}:\mathrm{R}^{n}_{+} \rightarrow \mathrm{R}^{m}_{+}$ *defined as (4.28) represents an efficient VRS technology and serves as a lower bound of any other efficient VRS mapping* $f:\mathrm{R}^{n}_{+} \rightarrow \mathrm{R}^{m}_{+}$.

Proof. Assume there is an arbitrary production correspondence **f()** satisfying the axioms L.1-L.2, and L.4 - L.7. Further, the origin is excluded from the graph **L(u)**, so **f** is concave on **D(X)** only, thus

$$\Sigma \lambda^{r} f(x^{r}) \leq f(\Sigma \lambda^{r} x^{r}), \ \Sigma \lambda^{r} = 1, \quad \textbf{(4.29)}$$
$$\lambda^{r} \geq 0, \ r = 1, \ldots, k$$

From the non-decreasing property of **f**, we have:

$$f(x) \geq f(\Sigma\lambda^r x^r),\ if\ x \geq \Sigma\lambda^r x^r, \qquad \lambda^r \geq 0,\ r = 1,\dots,k \tag{4.30}$$

The above relations imply

$$f(x) \geq \Sigma\lambda^r f(x^r),\ if\ x \geq \Sigma\lambda^r x^r, \qquad \lambda^r \geq 0,\ r = 1,\dots,k \tag{4.31}$$

Rewriting the above in another form, we have

$$f(x) \geq \{\ \Sigma\lambda^r f(x^r),\ if\ x \geq \Sigma\lambda^r x^r \qquad \Sigma\lambda^r = 1,\ \ \lambda^r \geq 0,\ r = 1,\dots,k\ \} \tag{4.32}$$

in general, and

$$f(x) \geq \max_{\lambda}\{\ \Sigma\lambda^r f(x^r),\ if\ x \geq \Sigma\lambda^r x^r \qquad \Sigma\lambda^r = 1,\ \ \lambda^r \geq 0,\ r = 1,\dots,k\ \} \tag{4.33}$$

in particular. Again, as shown in the proof of the proposition **4.2.1** and **4.2.2**, the data consistency ensures that

$$
\begin{aligned}
f(x) \geq \max_{\lambda} \{ & \Sigma\lambda^{r} f(x^{r}),\ x \geq \Sigma\lambda^{r} x^{r} \\
& \Sigma\lambda^{r} = 1,\ \ \lambda^{r} \geq 0,\ r = 1,\ldots,k \} \\
\geq \max_{\lambda} \{ & \Sigma\lambda^{r} u^{r},\ x \geq \Sigma\lambda^{r} x^{r}, \\
& \Sigma\lambda^{r} = 1,\ \lambda \in R_{+}^{k} \}
\end{aligned}
\qquad \textbf{(4.34)}
$$

This proves the "inner envelope" property of **(4.28)** as a linear VRS production frontier.

Q.E.D

Here we should pointed out that the frontiers bounded by $\mathbf{f^{V}(x)}$ must be concave on $\mathbf{D(X)}$. Finally, the nature of optimal solution of **(4.28)** can be explored in a similar manner as in the CRS case.

The above three propositions reveal that the family of DDPF, in each returns to scale category, is bounded from below by a linear production frontier. The linear production frontiers are "underwear" or "inner envelopes" covering the data sets.

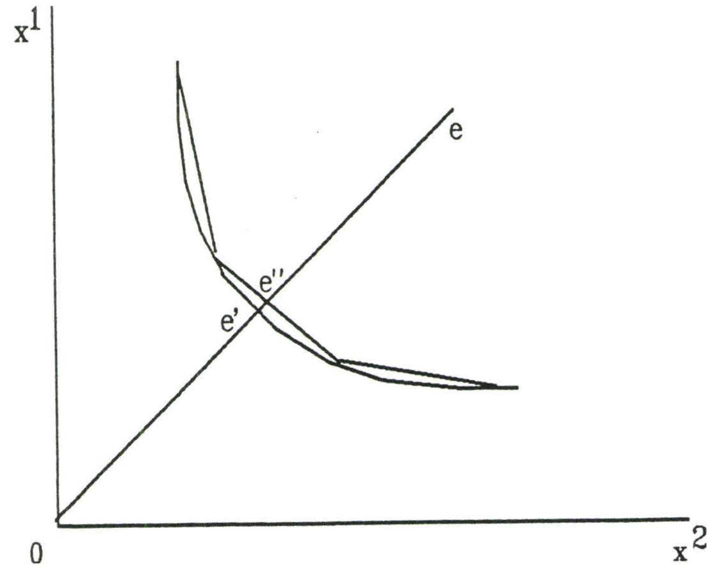

Figure 4.5 The "inner envelope" property and measurement of technical efficiency

This "inner envelope" property has an important implication for efficiency measurement. As discussed as in Chapter 2, the technical efficiency of a data point is measured by the radial distance from the data point to the corresponding point on a production frontier. In the input space illustrated as Figure 4.5, the technical efficiency measurement for an interior point **e** based on some **f(x)** is **oe'/oe**. Note **f(x)** has DDPF property in general. However, the propositions 4.2.1, 4.2.2, and 4.2.3 reveal that **f()** is bounded from below by a piece-wise linear frontier. Correspondingly, the measurement of technical inefficiency based on the linear frontier is **oe"/oe**. Apparently, no other gauge of technical efficiency yields a larger magnitude than this measurement.[11] Therefore, the measurement based on the linear frontier serves as an upper bound of any other measurement. This leads to a proposition regarding measurement of technical inefficiency:

> **Proposition 4.2.4** *Technical efficiency measurement based upon the linear frontier such as (4.1), (4.18) and (4.28) is the upper bound for technical efficiency measurement based upon any other DDPF.*

Proof of this proposition is omitted due to the apparent simplicity.

Finally, it should be noted that the propositions hold only as long as no random noise is present in the production process. If random noise is present, we would not be able to locate the linear production frontiers exactly and have precise efficeincy measurement.

---

[11] Otherwise, production possibility set would not be a convex set.

4.3. The Formulation of the Non-parametric Programming Approach for Efficiency Estimation

Based upon the propositions stated previously, in this and the next section we will examine the relationship between alternative models of technical efficiency measurement in the context of non-parametric programming approach.[12]

Suppose that the **CRS** frontier introduced in the last section is to be referenced. A linear programming problem follows immediately:

$$
\begin{aligned}
& \tau^c = \min \tau \\
& s.t. \sum_{r=1}^{k} \lambda^r x_i^r \le \tau x_i^* && i = 1,\dots,n \\
& \sum_{r=1}^{k} \lambda^r u_j^r \ge u_j^* && j = 1,\dots,m \qquad (4.35) \\
& \lambda^r \ge 0 && r = 1,\dots,k
\end{aligned}
$$

The objective value of the LP solution serves as an estimate of technical efficiency for the observation $(\mathbf{x}^*,\mathbf{u}^*)$ in input

---

[12] Recently, Seiford and Thrall (1990) examined the relationship between different models based on a numerical analysis. However, the analysis in this section follows the Lagrangian approach as employed in the exposition of optimal status in the Proposition 4.2.1.

space.[13] The choice variable λ's are defined as the intensity vector as before.

A counterpart measurement in output space can be obtained by solving

$$\begin{aligned} \tau^c &= \max \tau \\ s.t. \sum_{r=1}^{k} \lambda^r u_j^r &\geq \tau u_j^* \qquad i = 1,\ldots,n \\ \sum_{r=1}^{k} \lambda^r x_i^r &\leq x_i^* \qquad j = 1\ldots,m \\ \lambda^r &\geq 0 \qquad r = 1,\ldots,k \end{aligned} \tag{4.36}$$

Suppose the constraints are active for some optimal choice λ, then, obviously the objective value can be read as the ratio of the observed output $\mathbf{u}^*$ and a point on the imaged frontier. In the following presentation, however, we will concentrate on input space. The output counterpart of the measurement can be inferred in an parallel way.[14]

As explained in section 4.2, by imposing a further restriction on the intensity vector, we have the LP problem

---

[13] Note that this LP problem provides an estimate of efficiency for only one point (x^*,u^*) in the data set. A separate LP is required for each data point to obtain an estimate of its efficiency.

[14] Intuitively, efficiency can be measured either in terms of how many fewer inputs are required for a given output, or, how much extra outputs could be produced from the given inputs.

$$
\begin{aligned}
&\tau^{n} = \min \tau \\
&s.t. \sum_{r=1}^{k} \lambda^{r} x_{i}^{r} \leq \tau x_{i}^{*} && i = 1,\ldots,n \\
&\sum_{r=1}^{k} \lambda^{r} u_{j}^{r} \geq u_{j}^{*} && j = 1,\ldots,m \qquad (4.37) \\
&\sum_{r=1}^{k} \lambda^{r} \leq 1 && r = 1,\ldots,k
\end{aligned}
$$

which measures the efficiency under the restriction of NIRS.

In the case of variable returns to scale (VRS), the inequality of the intensity vector constraint is replaced by an equality constraint. Thus, the LP problem formulation becomes

$$
\begin{aligned}
&\tau^{n} = \min \tau \\
&s.t. \sum_{r=1}^{k} \lambda^{r} x_{i}^{r} \leq \tau x_{i}^{*} && i = 1,\ldots,n \\
&\sum_{r=1}^{k} \lambda^{r} u_{j}^{r} \geq u_{j}^{*} && j = 1,\ldots,m \qquad (4.38) \\
&\sum_{r=1}^{k} \lambda^{r} = 1 && r = 1,\ldots,k
\end{aligned}
$$

The formulation for measurements in the NIRS and VRS cases in output space can be derived parallel to equation (4.36).

As one may notice, this measuring scheme conforms with the theoretical efficiency structure described in Chapter 2. Formation of a piece-wise production frontier is the starting point to carry out a performance evaluation. The economic underpinning for the formulation can be easily perceived.

This approach has been closely followed by theorists, e.g. Afriat (1972) and R. Fare (1988).

On the other hand, management scientists evaluate a firm's performance by looking at a maximized ratio of its weighted outputs to weighted inputs, and subject to the condition that the similar ratio of every other firms be less or equal to unit. [Charnes et. al. (1978)] The data envelopment analysis (DEA) is devised to solve the non-linear programming problem

$$\max \gamma = \frac{\sum_{j=1}^{m} \alpha_j u_j^*}{\sum_{i=1}^{n} \beta_i x_i^*}$$

$$s.t.: \frac{\sum_{j=1}^{m} \alpha_j u_j^r}{\sum_{i=1}^{n} \beta_i u_i^r} \leq 1 \qquad \begin{array}{l} i = 1,\ldots,n \\ j = 1,\ldots,m \\ r = 1,\ldots,k \end{array} \tag{4.39}$$

$$\alpha_j, \beta_i \geq 0$$

to measure the technical efficiency for the observation $(\mathbf{x}^*, \mathbf{u}^*)$. The problem can be rewritten in a linear form as

$$\max. \gamma' = \sum_{j=1}^{m} \alpha_j u_j^*$$

$$s.t.: \sum_{j=1}^{m} \alpha_j u_j^r - \sum_{i=1}^{n} \beta_i x_i^r \leq 0 \qquad j = 1,\ldots,m \tag{4.40}$$

$$\sum_{i=1}^{n} \beta_i x_i^* = 1 \qquad i = 1,\ldots,n$$

$$\alpha_j, \beta_i \geq 0 \qquad r = 1,\ldots,r$$

The DEA model in this original form is capable only of handling CRS technology. In 1984, Banker, Charnes and Cooper (BCC) introduced a major improvement to DEA in order to capture alternative technologies characterizing production surface. The problem is formulated as:

$$
\begin{aligned}
& \max.\ \gamma' = \sum_{j=1}^{m} \alpha_j u_j^* - t \\
& s.t.: \sum_{j=1}^{m} \alpha_j u_j^r - \sum_{i=1}^{n} \beta_i x_i^r - t \le 0 \qquad j = 1,\ldots,m \\
& \qquad \sum_{i=1}^{n} \beta_i x_i^* = 1 \qquad\qquad\qquad\qquad i = 1,\ldots,n \\
& \qquad \alpha_j\ ,\ \beta_i \ge 0 \qquad\qquad\qquad\qquad\quad r = 1,\ldots,r
\end{aligned}
\tag{4.41}
$$

According to BCC, the unsigned variable **t** is introduced to capture the returns to scale effect. Increasing, constant or decreasing returns to scale are implied according to whether **t** is less, equal or greater than zero. Therefore, the equation **(4.41)** is a generalized DEA model. The equation **(4.40)** is regarded a special case of **(4.41)** with the term **t** set to zero (CRS technology).

There exists apparently two sets of formulation. The equations **(4.35)**, **(4,37)** and **(4.38)** are in the first set and the equation **(4.41)** in the second. The relationship and the consistency between the two approaches are now considered.

4.4 The Duality Relationship Between the Two Lines of Model Formulation

To reveal the duality relationship, we rewrite (4.38), the VRS formulation, as

$$\begin{aligned}
&\tau^{v} = \min \tau \\
s.t.\ &\sum_{r=1}^{k} \lambda^{r} x_i^{r} \le \tau x_i^{*} && i = 1,\ldots,n \\
&\sum_{r=1}^{k} \lambda^{r} u_j^{r} \ge u_j^{*} && j = 1,\ldots,m \\
&\sum_{r=1}^{k} \lambda^{r} \ge 1 && r = 1,\ldots,k \\
&\sum_{r=1}^{k} \lambda^{r} \le 1
\end{aligned} \tag{4.42}$$

note that the constraint on the intensity vector is decomposed into two separate parts.

This LP problem can be rewritten in a Lagrangian form as:

$$\begin{aligned}
\underset{\lambda}{Min}\ \underset{\theta}{\max}\ L(\lambda,\theta) = \tau &+ \sum_{i}^{n} \beta_i \Big(\sum_{r=1}^{k} \lambda^{r} x_i^{r} - \tau x_i^{*} \Big) \\
&+ \sum_{j}^{m} \alpha_j \Big(u_j^{*} - \sum_{r=1}^{k} \lambda^{r} u_j^{r} \Big) \\
&+ t_1 \Big(\sum_{r=1}^{k} \lambda^{r} - 1 \Big) + t_2 \Big(1 - \sum_{r=1}^{k} \lambda^{r} \Big)
\end{aligned} \tag{4.43}$$

where

$$\theta = (\beta_1, \ldots, \beta_m, \alpha_1, \ldots, \alpha_n, t_1, t_2)$$

is the vector of Lagrangian multipliers. Rearranging it, we have

$$\begin{aligned}\min_{\lambda} \max_{\theta} L(\lambda, \theta) &= \sum_{j}^{m} \alpha_j u_j^* - t \\ &+ \sum_{r}^{k} \lambda^r \left(- \sum_{j}^{m} \alpha_j u_j^r + \sum_{i}^{n} \beta_i x_i^r + t \right) \\ &+ \tau \left(1 - \sum_{i}^{n} \beta^r x_i^r \right)\end{aligned} \qquad \textbf{(4.44)}$$

where $\mathbf{t = t_1 - t_2}$ is an unsigned variable handling the effect of the constraints on the intensity vector. This form then leads to the dual formulation:

$$\begin{aligned}\max.\ \tau &= \sum_{j=1}^{m} \alpha_j u_j^* - t \\ s.t.:\ & \sum_{j=1}^{m} \alpha_j u_j^r - \sum_{i=1}^{n} \beta_i x_i^r - t \le 0 \qquad j = 1, \ldots, m \\ & \sum_{i=1}^{n} \beta_i x_i^* = 1 \qquad i = 1, \ldots, n \\ & \alpha_j\ ,\ \beta_i \ge 0 \qquad r = 1, \ldots, k\end{aligned} \qquad \textbf{(4.45)}$$

This is exactly the BBC formulation.[15]

---

[15] Note that the second constraint in (4.45) has an equality sign. This is because the Lagrangian multiplier in the last term of equation (4.44) is positive. In fact the multiplier is the measure of efficiency, which has an effective range (0,1].

The treatment of constraints apparently plays a pivotal role in equalizing the two approaches. As revealed by the duality relationship, the term **t** in DEA formulation corresponds to the constraints on the intensity vector in the primal formulation. The exact correspondence according to the global property can be checked by the following chart:

	Primal		*Dual*
CRS:	$\lambda^r \geq 0$	$\leftrightarrow$	$t = 0$
NIRS:	$\Sigma\lambda^r \leq 1$	$\leftrightarrow$	$t \geq 0$
VRS:	$\Sigma\lambda^r = 1$	$\leftrightarrow$	*t unsigend*

However, in the BCC formulation, **t** is an unsigned variable which is subject to LP evaluation. Therefore the BCC formulation in fact corresponds to the VRS version of the primal.

To see this correspondence, it should be noted that for each observation **(x,u)** the suggested LP problem yields a particular facet of the production frontier. Consequently, the entire production frontier is formed by applying the LP procedure repeatedly to all **k** observations. In the primal formulation, the stated property (e.g. CRS) is consistently enforced by imposing the corresponding constraint as each of the k facets of the frontier is built. As a result, the derived frontier maintains a global consistency. This consistency may still be preserved if the same constraint is consistently enforced in the dual problem. However, the BCC formulation does not maintain this consistency as **t** is unsigned. As a result, the entire frontier implied by **(4.45)** does not have a unified global property. In other words, VRS is its default global property.

Therefore, we conclude that "returns to scale" has a different underpinning in these two approaches. In the primal formulation, it refers to a global property whereas in the BCC formulation, it suggest a local property.

The fundamental discrepancy between the two lines of formulation can be traced to the level of basic axioms. BCC postulated the following convexity axiom for the production possibility set T:[16]

> *If* $(\mathbf{x}_r, \mathbf{u}_r) \in \mathbf{T}$, $\mathbf{r} = 1,\ldots,\mathbf{k}$, *and* $\lambda_r \geq \mathbf{0}$ *are non-negative scalars such that* $\sum \lambda_r = \mathbf{1}$, *then* $(\sum \lambda_r \mathbf{x}_r, \sum \lambda_r \mathbf{u}_r) \in \mathbf{T}$.

This condition is sufficient to ensure that we can build convex cones $\sum \lambda_r \mathbf{x}_r$ and $\sum \lambda_r \mathbf{u}_r$ in input and output space. However, it does not imply that the set T (or any graph linking x and u) a convex graph. Therefore, this condition is

[16] See Banker, Charnes and Cooper (1984), 1081. The notations are changed for the sake of presentation consistency.

less regulative than axiom L.7 introduced in Chapter 2.[17] Empirically, by this definition, the origin is excluded from the graph of input (output) mapping. Suppose there is some point to point mapping **f** transforming inputs to outputs, than, the above axiom implies that **f(0) = 0** is not regarded as (an implicit) data point, **f** is concave on, and only on, the domain **D(X)**. This is exactly the VRS case handled by proposition 4.2.3.

In Figure 4.6, there are four observations **a-b-c-d** in a data set. In our primal formulation, different constraints on the intensity vector alters our proposed frontier. In the CRS case, this proposed frontier is **o-b-c** while in the NIRS case it is **o-b-c-d**. In the VRS version of the primal and the BCC formulation, the frontier is concave only on $\mathbf{x}_a$-$\mathbf{x}_d$ and we have a frontier **a-b-c-d**. If section **a-b** (a facet) is referenced, we have IRS local property; if **b-c** is referenced, then CRS is obtained; if **c-d** is referenced, the DRS is noted. Having this discrepancy in mind, one should be warned that the efficiency gauged by the BCC formulation is equivalent to the standard measurement under VRS restriction. For the observation **f** in the figure 4.6, this is the distance **f-e'**.

This discrepancy leads to an alternative way of

---

[17] However, if axiom L.7 hold, above condition is implied.

interpreting technical inefficiency. The concept of "most productive scale size" (m.p.s.s) introduced by BCC refers to the frontier that is derived by imposing **t = 0**. Clearly the derived frontier is the CRS frontier (**o-b-c**) in the primal formulation. Total technical inefficiency (defined as **f-e"**) can be decomposed into two parts: pure technical inefficiency **f-e'** and scale inefficiency **e'-e"**. Even for an observation located on the portion **a-b** (or **c-d**), it may not be efficient since it may not reach (or exceed) the most productive scale size of operation, namely, **b-c**.

To conclude this section, we say that the two versions of non-parametric efficiency estimation approach is exactly matched in form of primal and dual if and only if the restrictions on the LP problem are consistently specified.

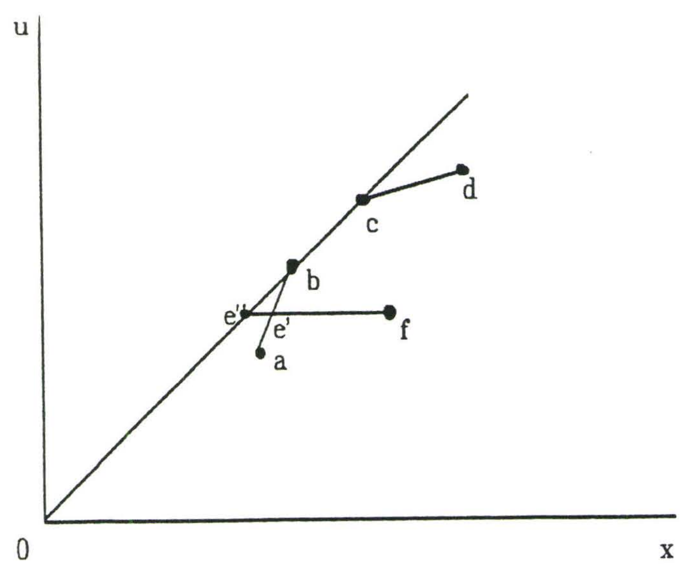

Figure 4.6 Returns to scale and m.p.s.s

4.5 Summary

In this chapter, we have reviewed one of most important approaches of efficiency estimation: the non-parametric programming approach. We checked the theoretical foundations of the non-parametric approach and alternative formulations of efficiency estimation. Also, we investigated the duality relationship between the two, with returns to scale as a key consideration. We demonstrated that the basic axiomatic assumption leads to different strategies in addressing the estimation problem. In the next chapter, we turn to the econometric approach to technical efficiency estimation.

CHAPTER 5. CONDITIONAL ESTIMATION OF TECHNICAL EFFICIENCY: A STOCHASTIC FRONTIER MODEL

In the realm of conventional econometric studies of the production process, the stochastic production frontier models bear some unique advantages in handling technical efficiency. This chapter concentrates on the stochastic frontier approach to technical efficiency estimation. In the first section of this chapter, we review alternative model specifications and estimation techniques. Based on this review, we propose a conditional estimation procedure to extract a measure of technical efficiency from a composite error structure.[1] The model employed is a generalized version of the stochastic frontier model, which assumes that the efficiency profile of a given set of production data has truncated normal distribution. As a result of the generalization, the capacity of stochastic model can be enhanced to cover wider spectrum of assorted efficiency profiles. The third section discusses the application of the model and estimation techniques.

---

[1] The section is based on a working paper (1990) by the author, which proposed conditional efficiency estimation procedures for both the truncated-normal/normal and the Laplace (double exponential)/normal error structure.

5.1 Stochastic frontier models: technique and Specifications.

5.1.1 The methodology.

The stochastic frontier (SF) model assumes that the mapping of an input set to output space is subject to the influence of some random disturbing force. Due to the presence of the random factor, the input and output are related in following functional form:

$$u = f(x) \cdot e^{v} \qquad -\infty < v < \infty \tag{5.1}$$

where **f()** is a deterministic kernel mapping the input set to the output space. The variable **v** is a random component which is unconstrained in sign and is assumed to be independent of input variables. Because of the stochastic component, the projected image of the input set on output space also has a probability distribution.

The effect of technical efficiency is handled by another variable in the SF model. To express the model in a conventional way, the production process has the following functional form:

$$\begin{aligned} u &= f(x)\ e^{\varepsilon} \\ \varepsilon &= v - t \qquad -\infty < v < \infty,\quad t \geq 0 \end{aligned} \tag{5.2}$$

where in the composite error structure **exp(v-t), t** is a non-negative variable representing technical inefficiency in production. The term **exp(-t)** in the error structure can serve as a standard measure of technical efficiency along the radial direction, ranging from zero to unity.

This composite error structure has some advantages over other models. In a conventional statistical model, if a production function is forced to maintain data consistency and a frontier property, an extraordinary outlier (due to some random shock) may result in the misplacement of the frontier and put the effect on the entire efficiency estimation. However, in a SF model, the modelling variable **v** serves as a filter screening out the effect of random shocks. An extraordinary outlier then has only a limited effect on the placement of the frontier. Figure 5.1 is an illustration of the SF model's production and efficiency structure. The observed data set is explained by a conventional parametric function **f()** which covers the mass of observations (snapshot A) . Further, the observations lying below **f()** are assumed to be distributed as $p_t(t)$ due to the influence of inefficiency factors (snapshot B) and the random noise **v** distributes as $p_v(v)$ (snapshot C).

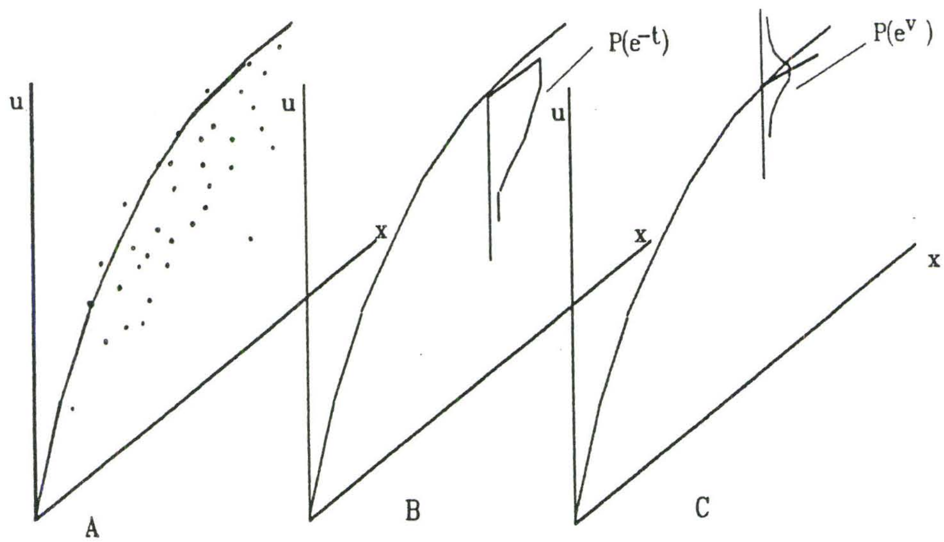

Figure 5.1 The structure of stochastic frontier model

5.1.2 Estimation Procedure of SF approach.

As we briefly outlined in Chapter 3, there are two successive steps in the entire estimation procedure; the first step pertains to production frontier estimation and the second one is the efficiency estimation.

In the first step, the maximum likelihood (ML) method is employed to estimate the production frontier. To obtain the ML function, the distributions $p_v(v)$ and $p_t(t)$ have to be pre-specified. Based on these specifications, the likelihood function can be expressed as:

$$L(\varepsilon_i;\beta,\theta_t,\theta_v) = \prod p_\varepsilon(\varepsilon_i;\beta,\theta_t,\theta_v) \qquad \textbf{(5.3)}$$

where β and θ are parameter sets to be estimated in f(x),

$p_v(v)$ and $p_t(t)$. Within the ML function, it should be noted that:

$$\varepsilon_i = \ln u^r - \ln f(x^r)$$

and if, as we will assume throughout, v and t are independent, we have:

$$p_\varepsilon(\varepsilon) = \int_0^\infty p_{(t,\varepsilon)}(t,\varepsilon)\ dt$$
$$= \int_0^\infty p_t(t)\ p_v(v)\,|_{v=\varepsilon+t}\ dt$$

Based on the estimates of the parameter sets, the technical inefficiency of an observation can be estimated in the second step [See Jondrow et. al. (1982)]. In this step, the distribution of **t** conditional on the composite error **(v-t)** is derived based on the Baysian rule, i.e :

$$p_t(t \mid \varepsilon) = \frac{p_{(t,\varepsilon)}(\varepsilon, t)}{p_\varepsilon(\varepsilon)}$$

Therefore, the first order moment of the origin for the distribution

$$E(t \mid \varepsilon) = \int_0^\infty t\, p_t(t \mid \varepsilon)\ dt \qquad \textbf{(5.4)}$$

and the mode of the distribution

$$M(t \mid \varepsilon) = \begin{cases} t \mid_{\frac{\partial}{\partial t} p_t(t \mid \varepsilon) = 0}, & t \mid_{\frac{\partial}{\partial t} p_t(t \mid \varepsilon) = 0} \geq 0 \\ 0, & t \mid_{\frac{\partial}{\partial t} p_t(t \mid \varepsilon) = 0} < 0 \end{cases} \tag{5.5}$$

may serve as the two alternative measures of technical inefficiency.

These two measures for each observations are conditional on specific composite error ε for each observations. The composite error ε is replaced by its estimated value in empirical work.

5.1.3 Alternative Model Specifications

Clearly, specification on the distribution $p_t(t)$ and $p_v(v)$ are crucially important in the SF model. Since it was introduced by Aigner, Lovell and Schmidt (1977), the SF model has had very few variations. Within the composite error structure, the pure random term **v** has been consistently (and reasonably) assumed to have a normal distribution. On the contrary, the distribution form for the technical efficiency term **t** has a few variations. The half-normal and exponential distributions are the most popular specifications for the efficiency term due to the simple distribution structure.

It should be noted that the distribution of the exponential of the term **-t**, (i.e., **exp(-t)**) describes the efficiency profile of a set of observations, say, an industry. In this regard, a flexible distribution of **t** implies a greater capacity to resemble the real world. To achieve the flexibility, Greene (1990) recently formulated the conditional estimation procedure for a Gamma-distributed stochastic frontier model. However, in his somewhat delicate model, the conditional estimation has no closed form and hence results in considerable complexity in estimation.

In the next section, we introduce a generalized model which assumes that the efficiency term **t** has a truncated normal distribution.[2] The advantage of this model, compared with the half-normal and the exponential models, is that a mode (or peak) is allowed in the normalized domain of technical inefficiency **(0, 1]**. It should be mentioned that the original model's structure is due to Stevenson (1980). The new development introduced in the next section is the conditional estimation procedure for the conditional mean and conditional mode.

5.2. Conditional Estimates of Generalized Models.

5.2.1 The Production Functions and the Error Structures

---

[2] Since this thesis was written, we have found a paper by Battese and Coelli (1988), which derived similar result in the context of panel data. However, they did not provide the mode for the conditional distribution. Beside, Greene (1990) gave a same result, though there is a sign error in the result.

Assume a production function has the following structure:

$$u = f(x)\,e^{\varepsilon} \qquad \textbf{(5.6)}$$
$$\varepsilon = v - t \qquad t \geq 0,$$

where $\mathbf{u} \in \mathbf{R}_+$ is output and $\mathbf{x} \in \mathbf{R}^n_+$ is an input vector. Within the error structure ε, $\mathbf{t}$ is a term reflecting technical inefficiency and $\mathbf{v}$ is a pure stochastic variable.

We make the following statistical assumptions:

A.1. *The pure stochastic variable* $\mathbf{v}$ *is distributed as* $\mathbf{N(0,\sigma_v^2)}$.

A.2. $\mathbf{t}$ *is a variable with a generalized truncated normal (GTN) distribution.* [3]

A.3. $\mathbf{v}$ *and* $\mathbf{t}$ *are statistically independent.*

A.4. *Both* $\mathbf{t}$ *and* $\mathbf{v}$ *are statistically independent of the input vector* $\mathbf{x}$.

In the Cobb-Douglas case, the above assumptions imply that

---

[3] The truncation point is set at zero and only the portion distribution for positive t is preserved.

$$Max: \prod_{r=1}^{k} p(\varepsilon^{r}) \tag{5.7}$$

$$\varepsilon^{r} = \ln u^{r} - \ln A - \sum_{i}^{n} \beta_{i} \ln x_{i}^{r}, \qquad i = 1,\ldots,n,$$

may serve as the objective function for maximum likelihood estimation.[4]

The probability density function (*pdf*) of GTN is[5]:

$$p_t(t) = \frac{1}{\sigma_t \sqrt{2\pi}\ [1 - F^*(-\frac{\mu}{\sigma_t})]} e^{-\frac{(t-\mu)^2}{2\sigma_t^2}}, \quad t \geq 0, \tag{5.8}$$

where $\mathbf{F*}(-\mu/\sigma_t)$ is a standard cumulative normal density function evaluated at $-\mu/\sigma_t$ and μ is the mean for the untruncated normal *pdf*.

By assumption **A.3**, and substituting out **v** by ε, we obtain the joint density function

$$p_{(\varepsilon,t)}(\varepsilon, \cdot t) = p(v\ ,\ t)\,|_{v=\varepsilon+t} \tag{5.9}$$

---

[4] See Stevenson (1980).

[5] See K.Bury (1975) p.154-155.

which can be integrated to a marginal density function

$$p(\varepsilon) = \int_0^\infty p(\varepsilon, t)\, dt$$

$$= \frac{f^*\left(\frac{\varepsilon + \mu}{\sqrt{\sigma_t^2 + \sigma_v^2}}\right)}{\sqrt{\sigma_t^2 + \sigma_v^2}\,\left[F^*\left(\frac{\mu}{\sigma_t}\right)\right]} \left[1 - F^*\left(\frac{\varepsilon\,\sigma_t^2 - \mu\,\sigma_v^2}{\sigma_v\,\sigma_t\,\sqrt{\sigma_t^2+\sigma_v^2}}\right)\right]$$

where $f^*()$ is a standard normal *pdf*. The mean and the variance for the density function (5.10) are given by

$$E(\varepsilon) = -\mu - \sigma_t \frac{e^{-\frac{\mu^2}{2\sigma_t^2}}}{\sqrt{2\pi}\,F^*\left(\frac{\mu}{\sigma_t}\right)} \tag{5.11}$$

and

$$V(\varepsilon) = \sigma_v^2 + \sigma_t^2\left(1 - \frac{e^{-\frac{\mu^2}{\sigma_t^2}}}{2\,\pi\,F\left(\frac{\mu}{\sigma_t}\right)^2}\right) - \sigma_t\,\mu\,\frac{e^{-\frac{\mu^2}{2\sigma_t^2}}}{\sqrt{2\,\pi}\,F\left(\frac{\mu}{\sigma_t}\right)} \tag{5.12}$$

respectively[6].

---

[6] It is easy to prove that these first order moments are generalization of the results given by Weinstein (1962).

If $\mu = 0$, the marginal density function reduces to

$$p(\varepsilon : \mu = 0) = \frac{2\, e^{-\frac{\varepsilon^2}{2(\sigma_t^2 + \sigma_v^2)}}}{\sqrt{2\pi}\,\sqrt{\sigma_t^2 + \sigma_v^2}} \left[1 - F^*\left(\frac{\varepsilon\, \sigma_t}{\sigma_v \sqrt{\sigma_t^2 + \sigma_v^2}}\right)\right] \tag{5.13}$$

which is the result given by Aigner, Lovell and Schmidt (1977) (where **t** has half-normal distribution).

The Baysian rule implies that the density function of **t** conditional on ε can be derived from the ratio $p(\varepsilon,t)$ to $p(\varepsilon)$, i.e.

$$p(t \mid \varepsilon) = \left\{\frac{1}{\left[1-F^*(-\mu/\sigma_t)\right]\, \sigma_t\, \sigma_v (2\pi)}\, e^{-\left[\frac{(t-\mu)^2}{2\sigma_t^2} + \frac{(\varepsilon+t)^2}{2\sigma_v^2}\right]}\right\}$$
$$\left\{\frac{f^*\left(\frac{\varepsilon + \mu}{\sqrt{\sigma_t^2 + \sigma_v^2}}\right)}{\sqrt{\sigma_t^2 + \sigma_v^2}\; F^*(\mu/\sigma_t)}\left[1 - F^*\left(\frac{\varepsilon\sigma_t^2 - \mu\sigma_v^2}{\sigma_v \sigma_t \sqrt{\sigma_t^2 + \sigma_v^2}}\right)\right]\right\}^{-1} \tag{5.14}$$

From the conditional density function, it follows immediately that:

Proposition 5.1. *If the term for technical efficiency* **t** *has a generalized truncated normal distribution and random variable v has a normal distribution, then the mathematical expectation of* **t** *conditional on* ε *is*

$$E(t \mid \varepsilon) = \frac{\sigma_t \sigma_v}{\sqrt{\sigma_t^2+\sigma_v^2}} \left[\frac{f^*(A)}{1-F^*(A)} - A \right] \tag{5.15}$$

where

$$A = \frac{\varepsilon \sigma_t^2 - \mu \sigma_v^2}{\sigma_t \sigma_v \sqrt{\sigma_t^2 + \sigma_v^2}} \tag{5.16}$$

proof. The result is obtained by integrating the conditional density function multiplied by t.

Further

Proposition 5.2 *The mode for the conditional distribution $M(t|\varepsilon)$ is given by*

$$M(t \mid \varepsilon) = \begin{cases} \dfrac{\mu\sigma_v^2 - \varepsilon\sigma_t^2}{\sigma_t^2 + \sigma_v^2} & \varepsilon \geq \mu\sigma_v^2/\sigma_t^2 \\ 0 & \varepsilon < \mu\sigma_v^2/\sigma_t^2 \end{cases} \tag{5.17}$$

proof. The result can be derived from the first order derivative of the conditional distribution.

It is easy to show that the results are identical to those derived by Jondrow et. al. (1982) if the distribution of **t** is truncated at $\mu = 0$.

5.3 Implications of the Model

As revealed in Chapter 3, there are many possible patterns for technical efficiency profile. The early versions of the SF model approximated these patterns by either a half-normal or exponential distribution. Correspondingly, the normalized efficiency index **exp(-t)** exists in the range (0,1] and the mode of the distribution is fixed at unity. [pattern A in Figure 5.2]. The limitation of these early versions is quite obvious. If the "true" pattern of the real world is a distribution with a mode located in the range (0,1], then neither of the previous versions can authentically approximate the real world.

The major achievement of the current generalization is obtained by introducing the location parameters μ (which was forced to be zero in earlier versions). Introducing the location parameter make it permissible for an efficiency distribution to have a mode within the range (0,1]. As a result, the risk of mis-interpretation of the real pattern (e.g., B in Figure 5.2) is substantially reduced. Nevertheless, there is a cost for this generalization: some extra effort has to be paid to handle the additional parameters and the estimation procedure becomes a more complicated task.[7]

---

[7] The gradient vector employed in maximum likelihood estimation is given in Appendix 1.

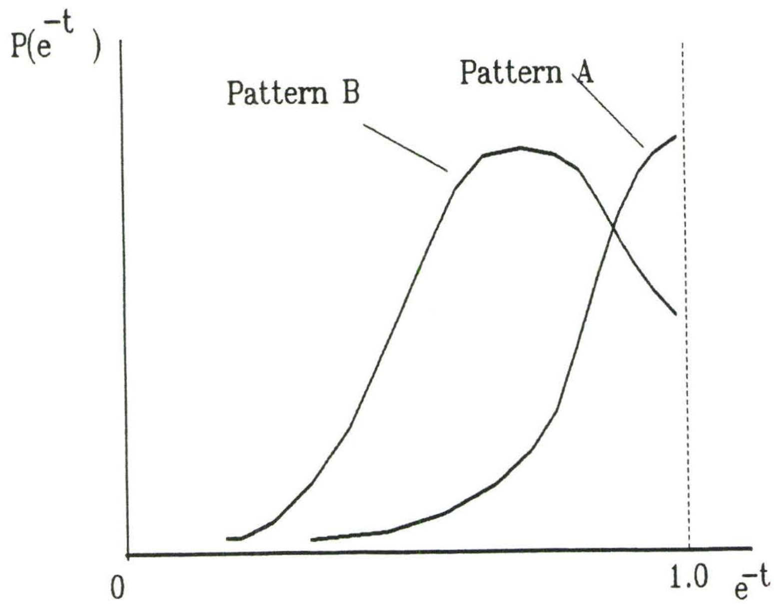

Figure 5.2 Two patterns of efficiency profiles

5.4 Summary.

The advantages derived from above generalization are obvious. The truncated-normal/normal error structure provide enhanced flexibility for the SF model. On one hand, production function estimation is benefits from a more adequately accommodated error structure while on the other hand, it is expected that the model can cover a wider spectrum of technical inefficiency profiles. However, extra effort may have to be paid to handle the delicate structure. Finally, one should understand that there is no guarantee that the efficiency term can be successfully extracted from the composite error. The reliability of estimation result should be verified.

CHAPTER 6. THE EXPERIMENT DESIGN FOR A SIMULATION STUDY

Making use of the knowledge presented previously, this and the next chapter are devoted to a simulation study designed to test the ability of the stochastic frontier (SF) and the DEA model to measure efficiency under assorted situations.[1] This chapter concentrates on the design of the simulation experiments and related issues.

In contrast to flourishing empirical applications, experimental investigation on alternative efficiency estimation methods is rarely explored. However, a few reported studies, each with a different focus, are worthy of note.

To test the SF model, Aigner, Lovell and Schmidt (ALS) (1977) designed a series of experiments to compare outcomes obtained from different specifications of the production and efficiency structure. Their focus is the performance of the SF model under different settings. Another early experimental study of the stochastic frontier model is due to Olson,

---

[1] For the sake of simplicity, SF and DEA are used to refereed to the stochastic frontier model and data envelopment analysis respectively in the rest part of the presentation.

Schmidt and Waldman (OSM) (1980). The major concern of their study is alternative estimation techniques rather than model specifications. In the study, the effectiveness of MLE, 2STEP (Two Step Newton-Raphson) and COLS (Corrected Ordinary Least Square) estimation techniques were assessed. Two points should be noted about these earlier studies. First, the emphasis of the experiment is the estimation of the production frontier. Technical efficiency estimation itself is not the focus. Second, these studies examine the competence of alternative model specifications or alternative estimation procedures with the same methodology. Little effort was made to compare effectiveness of alternative approaches (that is, across econometric and non-econometric approaches).

A comparative study of estimation techniques with more distinct methodologies was reported more recently. Banker, Charnes, Cooper and Maindiratta (BCCM) (1988) conducted a simulation study of alternative efficiency estimation schedules. They compared a DEA model with a conventional OLS model and assessed their performance. This study brought efficiency estimation into the focus and the models based on distinct methodologies are compared. BCCM report that the DEA model has an unambiguous superiority over the OLS in regards to the accuracy of efficiency estimation. However, because a somewhat naive OLS model was employed in the study, the question remains open regarding more recent estimation models.[2]

The present comparative study focuses on two efficiency estimation methods: the stochastic frontier model and the DEA model. Since our focus is the estimation of technical efficiency. Thus, we do not divert our attention to the analysis of the production structure.

This chapter proceeds as follows. Section one discusses the details of the proposed experiments. In section two, we consider performance assessment criteria. Section three discusses the data generation and the estimation techniques adopted in the experiment.

6.1 Assignment of Experiments

Before the discussion of experimental design, two notions we used in this and the next chapter should be explained.

Efficiency index

As demonstrated in Chapter 2 and 4, we know that Farrell's technical efficiency measurement (and the DEA estimation) has an effective range **(0,1]**. In SF model,

---

[2] The recent empirical comparative applications of parametric and non-parametric methods can be found in Ferrier and Lovell (1990) and Bjurek, Hjalmarsson and Forsund (1990).

however, efficiency term is represented by variable **t** $(t \geq 0)$. Therefore, we define

$$z = \exp(-t)$$

as a normalized efficiency index. Clearly, since **z** has an effective range **(0,1]**, thus this transformation provides the SF and DEA efficiency estimates with a comparable basis.

The Efficiency Profile

In our experiments, the notion of "efficiency profile" is refereed to the frequency distribution of **z**. Provide that **t** has a truncated normal distribution as **(5.8)**, the distribution of **z** given by:

$$p_z(z) = \frac{1}{z\,\sigma_t\sqrt{2\pi}\,[1 - F^*(-\frac{\mu}{\sigma_t})]}\, e^{-\frac{(\ln z + \mu)^2}{2\sigma_t^2}}, \quad t \geq 0, \qquad (6.1)$$

with the mean and the variance equal to:

$$E(z) = \mu_z$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_t\,F^*(\mu/\sigma_t)} \int_0^1 e^{-\frac{(\ln z + \mu)^2}{2\sigma_t^2}}\, dz \qquad \mathbf{(6.1.a)}$$

and

$$V(z) = \frac{1}{z\,\sigma_t\sqrt{2\pi}\,F^*(\mu/\sigma_t)} \int_0^1 (z - \mu_z)^2\, e^{-\frac{(\ln z + \mu)^2}{2\sigma_t^2}}\, dz \qquad (6.1.b)$$

Further, we emphasize that in our experiments, the comparison is made between the samples of the "true" efficiency indexes and their estimates. The parent or population distribution from which the efficiency indexes are drawn has only a secondary importance. Recovering a parent distribution is not our intention.

Now we proceed on to experiment design. The proposed experiments are grouped into three sets. In the first set, we evaluate the performances of DEA and SF under assorted efficiency profiles. In the second set of experiments, the two methods are applied to the production processes with different returns to scale and their performances are examined. The last set of experiments is designed to test the robustness of the two approaches regarding to various magnitudes of random noise. We discuss them in turn.

Experiment Set I: Exploring the Efficiency Profile

This experiment set is designed to examine the performance of both models under assorted underlying efficiency profiles. We consider four types of profiles:

<u>Case A</u>. A "J type" distribution (the right tail of a truncated normal distribution).

<u>Case B</u>. An efficiency profile from a typical truncated normal distribution (with a tail being truncated).

<u>Case C</u>. A profile with large variance.

<u>Case D</u>. A left-skewed efficiency profile.

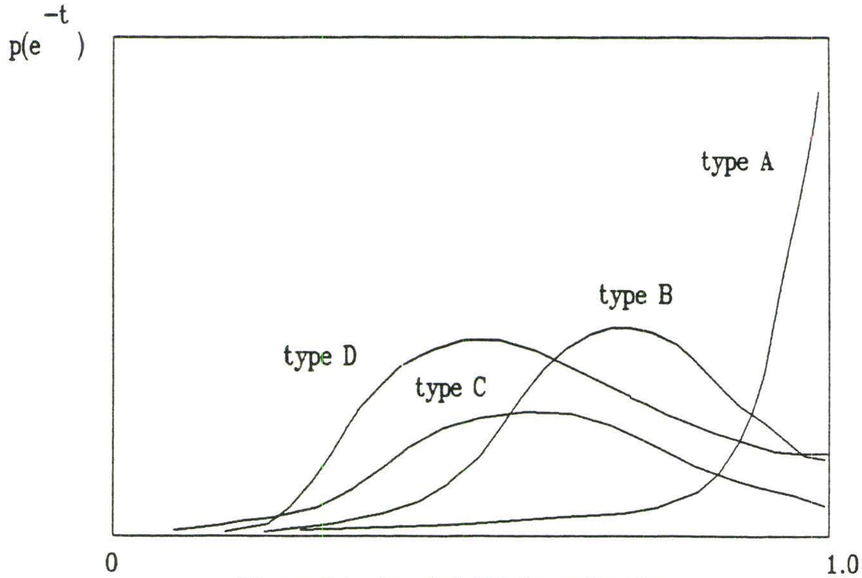

Figure 6.1. Assorted Efficiency Profiles

These four types illustrated in Figure 6.1 are designed to represent four possible efficiency profiles.[3] It is worth pointing out here, that types "A" and "B" are typical truncated normal distribution. The type "A" is in fact a tail of normal distribution while type "B" is a normal distribution with a tail being cut off. Further, types "C" and "D" follows no particular distribution. They are included to allow examination of the SF model when the assumed efficiency distribution (e.g. truncated-normal) is not in fact the true efficiency distribution. Admittedly, C and D are not wildly

---

[3] Though these patterns are hypothetical ones, some empirical evidences for these possible patterns can be found in Deprins, Simar and Tulkens (1985) and Bjurek, Hjalmarsson and Forsund (1990).

different from a truncated normal. However, they should be sufficiently different to test the robustness of the SF model.

In experiment set I, the production function used to generate the production data has a CRS property and is fixed throughout the experiment.

Experiment Set II: Exploring Returns to Scale Specification.

As discussed in Chapter 4, the returns to scale assumption affects the interpretation of a production process and the measurement of technical efficiency. The SF model and the DEA model are fundamentally different in the way they handle the returns to scale property. In the SF production model, the returns to scale property is revealed as an *ex post* outcome derived from parameter estimates. It is not necessary to specify the returns to scale property prior to estimation. Consequently, the risk of mis-specification (from this sources) usually does not exist.[4] On the other hand, mis-specification of returns to scale can cause problems in the DEA approach because a pre-specification of returns to scale property is a necessity. As shown in Chapter 4, three types of returns to scale frontiers, i.e., constant returns to scale (CRS), non-increasing returns to scale (NIRS) and variable

---

[4] There is indeed a possibility of having less accurate estimates of the returns to scale parameters. This, however, is not a specification error.

returns to scale (VRS), can be built for a given a data set. In fact, the VRS frontier nests below the NIRS frontier and both of them nest below the CRS frontier. In order to measure technical efficiency accurately, one is obligated to select the correct DEA model that is consistent with the real world. Otherwise, the efficiency estimation cannot be accurate. Nevertheless, our knowledge of the real world is limited and a specification error seems to be a persistent threat to the DEA approach.

Retaining the Cobb-Douglas production structure, our second set of experiments includes two hypothetical scenarios:[5]

Case A: The production function exhibits constant returns to scale (CRS).

Case B: The production function exhibits decreasing returns to scale (DRS).

Pretending that reality is unknown to us, three DEA models and one SF model will be applied in each case. It is expected that the consequence of mis-specification in DEA can be revealed. The experiments have a differet implication to

[5] With a Cobb-Douglas production structure, we may have one of three cases: constant returns to scale (CRS), decreasing return to scale (DRS) and increasing returns to scale (IRS). However, the IRS case has been ruled out since it violates over convexity axiom. Further, in the Cobb-Douglas function, only the DRS case is consistent with the NIRS assumption.

SF model. In many situations, the depressing effect of technical inefficiency on a production process is similar to the one of decreasing returns to scale. Therefore, the ability of SF model to separate these two effects is under examination. Table 6.1 is a panel showing the combinations of correctly and incorrectly specified situations for DEA models.

Table 6.1 Return to Scale Specification

DEA models applied:	CRS	NIRS	VRS
True property:			
CRS	Y	N	N
DRS	N	Y	N

Note: Y - correct specification.
N - incorrect specification.

In this and the next experiment set, the efficiency profile is chosen as Type B in the experiment set I. One should recall that this type of efficiency profile comes from a typical truncated normal distribution. Further, the results of experiment set I show that both models perform better with this profile. Thus the profile B is chosen as a default profile.

Experiment set III: Exploring the Sensitivity to the Disturbance of the Random Error.

This set of experiments is designed to inspect the robustness of both approaches with respect to the various levels of random noise. Generally, the SF model has a sophisticated design to filter out random noise and extract the efficiency factor from a mixed disturbance. The DEA model, on the other hand, less sophisticated in this regard. It does not provide any accommodation for random noise and, as a result, combines the effect with the inefficiency factor. However, with a louder noise, the robustness of the built-in filter in the SF model is also worthy of investigation. To extract systematic signals from a stream, heavily contaminated by larger noise, has some obvious difficulties. The SF model may fail to distinguish the efficiency factor from the random noise and to separate them effectively. It is unclear _a priori_ which model will perform better with larger noises.

In this experiment set, the standard deviation of the random noise component will be set at four levels so that the relative competence of both approach can be checked. In Case A, the standard deviation of the random noise is set at 0.01. In other words, this level on average is about 1 percent of full scale of the technical efficiency index, which has a range (0,1]. The noise level in Case B is the default noise level adopted in other sets of experiments. The standard deviation of the random noise at this level is 0.03. In Case C, the standard deviation is set at 0.06. Finally, the standard deviation of random noise is set at 0.09.

The assignments and the settings for the three experiment sets discussed above are summarized in Table 6.2. The table is read as follows: On the diagonal cells (boxed), one finds the variable factors being tested. The off diagonal cells may be viewed as control settings. Horizontally, in experiment set I, we investigate the effect of variations of the efficiency profile. The production function here has a CRS property and the noise level is set at a "moderate" level. On the second line, we have the settings of experiment set II that examines the effects of specification error with respect to returns to scale property. The efficiency profile in this case is fixed as type B and the noise level is set again at a moderate level. Finally, in experiment set III, we vary the random noise to different levels and examine the effects of such changes. The efficiency profile is set to be type B and the production function has a CRS property.

As indicated in the table, experiment set I has four cases, each corresponds to a different efficiency profile. Two basic cases in Experiment set II yield four possible situations: correctly specified and mis-specified situations with the CRS and the DRS situations. In addition to these, we add two experiments to examine the consequence of applying the VRS version of the DEA model to both of the CRS and DRS production data. Finally, experiment set III has four cases corresponding to various random noise levels.

For each case, both the DEA and the SF models are applied.

Table 6.2 Assignment of Experiment Sets

	Efficiency Profiles	Property of Production Function	Rel. random noise level (in std.)
Exp. set I :	Type A B,C and D	CRS	0.03
Exp. set II:	Type B	CRS, DRS	0.03
Exp. set III:	Type B	CRS	0.01, 0.03, 0.06 and 0.09

6.2. Performance Assessment Criteria

To assess the performance of the SF and the DEA estimation of technical efficiency, we implement a series of tests. The tests are grouped to cover three aspects of performance. The first group of tests focuses on the accuracy of individual efficiency index estimation. These tests attempt to measure how well we have done in estimating the efficiency of individual units. The second group of tests examines the ability of the two methods to rank the observations by efficiency. The third group of tests is designed to look at the ability of the methods to fit a given efficiency profile. In other words, it indicates how the distribution of the estimated technical efficiency resembles its original profile. Efficiency profile recovery can be viewed as a measure of performance of the model.

Test Group 1. Individual Index Estimation.

The first group of statistics concentrates on the accuracy of individual efficiency index estimation.

Average Deviation.

The average deviation is the average of the estimated technical efficiency index less the "true" value. The statistics reflect if a model over- or under-estimates the efficiency index. However, since the over-shooting and under-shooting values may offset each other, this statistics cannot tell the whole story.

To detect the sources where the deviations are generated, we divide the efficiency indexes and their estimates into three sub-groups according to the true index value: the first group contains the 25% most efficient observations; the second group contains the middle 50% observations; the last group contains the 25% least efficient observations. The average deviations for these three groups are reported individually. Since these grouped measures all contribute to the total, the total average deviation is in fact a weighted average of the three.

As well as showing where in the distribution of efficiency indexes, estimation is good and where poor, the separation into three groups might allow one to choose one method that is more accurate for the part of the distribution of most interest (e.g., very inefficient firms).

Average Absolute Deviation.

The average absolute deviation is the average of absolute difference between estimated technical efficiency index and the corresponding actual value. It indicates average estimation accuracy. A small average implies a better estimation quality. However, as an absolute value, it cannot indicate the direction of bias.

Test Group 2. Rank Information.

It is often claimed that technical efficiency is a comparative measurement of performance. In this regard, ranking may be a more effective indicator for a group of observations than the efficiency index itself. To check whether the alternative estimation procedures can provide correct ranking of observations, we chose Spearman's rank order correlation coefficient as a measuring gauge.

The theoretical definition of the coefficient is:

$$r_s = \frac{\sum(R_i^{est} - \overline{R}^{est})(R_i - \overline{R})}{\sqrt{\sum(R_i^{est} - \overline{R}^{est})^2 \sum(R_i - \overline{R})^2}} \tag{6.2}$$

where $\mathbf{R}_i$ and $\mathbf{R}_i^{est}$ are ranks of original and estimated efficiency index for observation **i** respectively and bars denote sample means of the ranks.[6]

Test Group 3. Distribution of Technical Inefficiency.

The frequency distribution of the technical inefficiency index for **n** observations is obtained by binning the **n** data points appropriately. To avoid information loss resulting from binning, we set the numbers of the bins to be 26 (besides the end bin, there are 25 bins running from 0.4 to 1.0 at width of 0.024). Further, we define $\mathbf{f}^{ori}$ and $\mathbf{f}^{est}$ as the binned frequency vectors for the original and the estimated efficiency indexes (from either DEA or SF). The i-th elements of the vectors are the number of the observations falling into the bin.

---

[6] If there is a tie among the ranks, it will be replaced by its mid-ranking value. Considering tied situation, we use the formula provided by Press, Flannery, Teukolsky and Vetterling (1989) to calculate Spearman's rank order correlation coefficient.

Chi-square Test

The Chi-square statistics is employed to test whether the distribution of the estimated efficiency indexes is significantly different from the original one. For the null hypothesis that the estimated distribution is identical with the original one, the Chi-square statistic is defined as:

$$\chi^2 = \sum_i \frac{(f_i^{est} - f_i^{ori})^2}{f_i^{est} + f_i^{ori}} \tag{6.4}$$

In this formula, both $\mathbf{f}^{ori}$ and $\mathbf{f}^{est}$ are now treated as samples. Thus the denominator is the sum, rather than average of the two terms.[7]

These test statistics reflect different aspects of estimation quality. There is clearly no unique measure of quality, but we think these measures capture the key features that might be of interest.

6.3 Other Specifications and Computational Techniques.

Data Generation and Sample Size

Three sets of variables, i.e. the production data, the

[7] For a discussion of the Chi-square statistics, see W.H. Press et. al., 488-489.

efficiency index, and the random noise component, were generated from independent sources as follows.

(1). Production data

For the sake of simplicity, we assume there are two inputs (L,K) and one output in a Cobb-Douglas production structure:

$$u = A\, L^{\alpha}\, K^{\beta}\, e^{v-t} \qquad \begin{array}{c} t \geq 0\ ,\ -\infty < v < \infty \\ 10 < L \leq 20,\ 20 < K \leq 30 \end{array} \qquad \textbf{(6.6)}$$

Data for both inputs L and K are generated from a uniform distribution with the range shown. The observable output **u** is then obtained from the function. The constant term is set to unity. The other parameter settings vary across experiments and will be given in the next chapter.

(2) The efficiency term

The efficiency term **t** is generated from a truncated normal distribution for the profiles "type A" and "B". The efficiency index **exp(-t)** then can be obtained correspondingly, However, in experiment set I, case C and D are designed to test the robustness of the two methods under some particular efficiency profiles. In these two cases, samples resembling the desired efficiency patterns as described in section 6.1 were created for the simulation study.

(3) The random variable

The samples of the random variable, **v**, are drawn from a normal distribution with a zero mean and a standard deviation that vary according to particular experiment set (See Table 6.2).

After the variables **t** and **v** are generated, they are merged with output **u** so that each sample contains only three series: two inputs and a "contaminated" output.

The sample size of experiment and number of replications are worth of discussion. As a general principle, larger the sample size and more replications, better the results. This is not only true for the SF model but also true for the DEA model since more data points yield a better approximation of the given production frontier from which data are generated. However, our experiments are confined by some inherent limitations. First, with a large sample size and many replications, computation burden for DEA could increase substantially. In the linear programming problem of DEA, the constraints increase as the sample size swells. As a result, the size of the tableau can be enlarged geometrically. This will reduce the computation speed siginificantly. Moreover, the DEA efficiency measurements for a given data set of size **n** is obtained by solving **n** LP problems repeatedly, each for a data point. This is another consideration to limit sample size and numbers of replication within a reasonable range. Therefore, we should emphasis that the results from our experiment pertain to finite sample properties of the methods under examination. The large sample properties are simply not our intention.

With these considerations, we set sample size at a moderate level with each sample contains 100 observations. However, for each case under examination, we draw 5 samples and only control the efficiency terms. Therefore, we have 5 replications in each case.[8] For instance, in case A of experiment set I, an identical series **t** is merged with five sets of input-output data and 5 drawings of random variable **v**, and five samples are generated. The estimation procedures are then implemented to all five samples yielding five sets of estimates.

Case A and B in the experiment I have a typical truncated normal distribution. In particular, the Case B is a norm for both experiment sets II and III. In these two cases (A and B), we add supplementary experiments. In each case, the five samples are generated in such way that one set of production data and pure noise is merged with five efficiency index vectors drawn from same population.

Normalization of the Efficiency Indexes

---

[8] Thus our experiments, designed in the last section, include 6800 LP problems, each with more than 100 constrants, to evaluate the technical efficiency for each data point.

Since the DEA method measures the comparative efficiency relative to a best performance observation at 100% efficiency, it seemed desirable to normalize the sample of efficiency indexes to put SF and DEA on an equal footing by dividing the sample indexes by the maximum of them. Thus the best performance observation in each data set has at least one efficiency index equal to one. It can be shown that this transformation does not affect the ML estimation for the SF model (See Appendix 2).

Summary of Experimental Results.

In most of the cases except case A and B in experiment set I, we will report only summarized results. The method of summarization will be addressed in next chapter.

Estimation Techniques

The LP programming formulations employed in the experiments are identical with those given in Chapter 4. Technical efficiency is measured along the direction of output axis. The computation procedure is carried out by a FORTRAN program interfacing with LINDO.[9] The program set up the data tableau for the LP problem and called LINDO to solve the problem iteratively. The FORTRAN code is provided in Appendix 3 to the thesis.

---

[9] LINDO (Linear, Interactive aNd Discrete Optimizer) was developed by LINDO systems at University of Chicago.

The SF model employed is the truncated-normal/normal model discussed in Chapter 5. The estimation is implemented by a FORTRAN program, which is a maximum likelihood function estimation procedure using Davidon-Fletcher-Powell (DFP) algorithm.[10] The program conducts a multi-dimension search for the maximum (minimum) of a (negative) likelihood function following the direction indicated by a Hessian matrix, which is constantly updated by the gradient information at run time. The gradient vector for the ML function is given in Appendix 1.[11] It should be noted that the same algorithm was employed by Meeusen and van den Broeck[12] (1977), Olson, Schmidt and Waldman[13] (1980) and Greene(1990) and proved to be an effective procedure.

The input data, the efficiency term **t** and the random variable **v** were generated by the random number generator supplied by MINITAB.[14] For the efficiency term **t**, a large

---

[10] Major reference for the algorithm can be found in Press et. al. (1988). Chapter 10.

[11] The code is provided in Appendix 4 to the thesis.

[12] The estimation package used was MINUIT, the method is incorporated into the package by J. James and M. Roos of CERN at Geneva.

[13] The estimation package is the Goldfeld-Quandt package developed by the Princeton Econometric Research Program.

[14] MINITAB is a statistical and data processing software developed by Minitab Inc.

number of observations (8000) were generated for given population parameters and then truncated. Afterward, samples of 100 were drawn from the truncated distribution.

The data editing and compiling jobs for the experiments were accomplished on a 80286 based PC computer and the major computation tasks were implemented on the VAX system of McMaster University.

CHAPTER 7. THE RESULTS OF SIMULATION STUDY

As explained in the last chapter, our comparative study of alternative efficiency estimation methods contains three sets of experiments. The results of the experiments are presented in this chapter. The chapter is arranged as follows: The first section assesses the performance of the SF and the DEA approaches under assorted efficiency profiles. The second section compares the robustness of the two approaches under alternative returns to scale scenarios. Section three examines the robustness of the two methods under alternative assumptions about the size of the random noise component. The final section provides a general assessment of the two efficiency measurement models.

7.1. Experiment Set I: Efficiency Profile

The four types of frequency distributions considered for the efficiency index are plotted in Figure 7.1. The four cases are chosen to provide a variety of testing frameworks for SF and DEA estimation.[1] Profiles A and B are the actual samples randomly drawn from a truncated normal distribution (with different shapes because of different truncation points) while profiles C and D are not. C is chosen to be somewhat flatter than a truncated normal while D is chosen to be left skewed. The truncated-normal (efficiency)/normal (random noise) model is a correct specification for the SF model discussed earlier only for cases A and B and the rest of the chapter (the second and third sections) focuses on these cases. Nevertheless, it seemed worthwhile exploring some cases where the efficiency error did not come from a truncated normal distribution. We discuss these four cases in turn.

Case A. Efficiency Profile Type A.

In this case, the mass of the distribution is close to unity and the distribution is the truncated tail of a normal distribution (with more than half of the distribution being cut off). In the experiments presented, we first combined one drawing of the efficiency index with 5 drawings of CRS production data and 5 drawings of random noise.[2] In this way,

---

[1] To justify the paramter settings for these profiles, we suggest that emperical investigations by Deprins et al. (1985) and, more recently, by Bjurek et al. (1990) should be referenced.

[2] The way in which capital, labour and the random noise component are generated is discussed in the previous chapter.

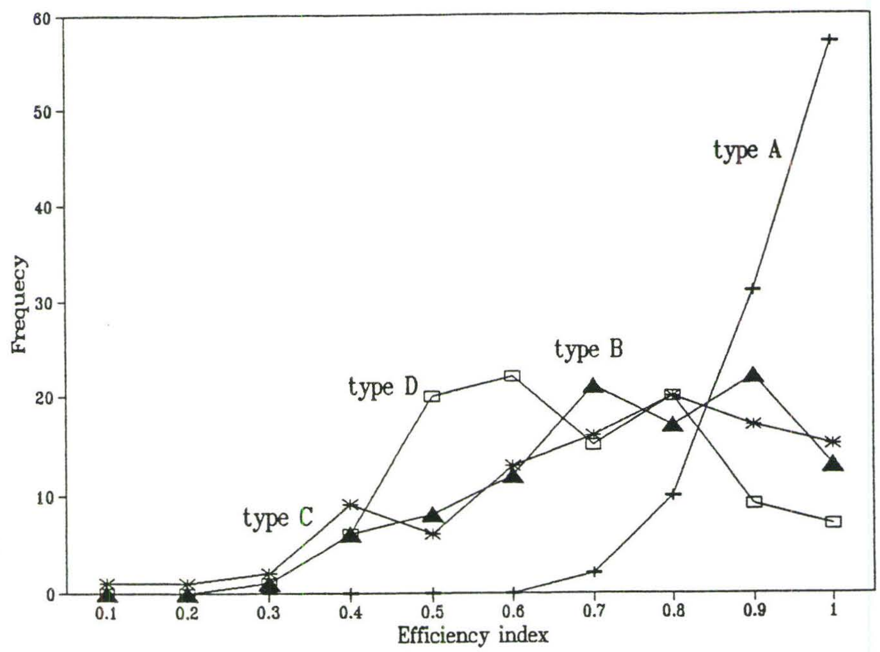

Figure 7.1 Distribution of Exp(−t)

five samples were generated. Both SF and DEA estimation procedures were then implemented. The theoretical (and population) parameters were used to serve as the initial values for the maximum likelihood (ML) estimation procedure. The purpose of this choice is to minimize computational error in the ML procedure.[3]

The results of these experiments are listed in Table

[3] It should be noted that the experimental data sets have been contaminated by random noise. Thus our choice just places the ML function value at some neighbourhood of the global maxima. The steps to reach the maximum vary from one to more than forty. However, a line minimization procedure is still involved in each step to search for the minimum along one direction. See Press et. al. (1988), 324-328.

7.1.A.1.[4] The statistics introduced in Chapter 6 for the five sample estimations are listed in the columns of the table. In the top section entitled "individual efficiency estimation", the first row reports the average deviation of the estimate from the true value (in particular, we report the average of the estimate minus the true value) over all 100 observations; the next three rows report the deviations in various sub-parts of the sample (from more to less efficient observations) while the fifth row reports the average absolute deviation over the 100 observations. The last two rows report the Spearman's rank correlation coefficient and the Chi-square value as explained in the last chapter. The last two columns of the table require further explanation. The first five rows (in these columns) report the averages across the 5 samples. However, no meaning can be attached to an average of the rank coefficients or of the Chi-square values. We report instead statistics for pooled estimation. In addition, the average absolute deviations for the pooled estimation are also reported. By pooled estimation we mean that we have taken the 5 separate estimates of the efficiency of each observation (or "firm") and formed an average estimate of the efficiency of

---

[4] The statistics for the efficiency index exp(-t) in the title of the table are population statistics. The formulas for converting the statistics for the variable t to those of exp(-t) are given in Chapter 6.

that observation.[5]

As a second part of this experiment, five drawings (sets) of efficiency indexes were merged with one set of the CRS production data and random noise[6]. The results are given in Table 7.1.A.2. As we have five drawings of efficiency indexes in this part of the experiment, pooled estimation is not applicable and hence not shown in the Table 7.1.A.2 (n.a.'s in the table).

The results of these two tables will be discussed together. First, for individual efficiency index estimation (the upper part of the tables), SF shows no significant tendency to over or underestimate the efficiency index. However, for the most efficient observations, it frequently undershoots the target.[7] On the other hand, DEA shows a consistent tendency to undershoot the target in all parts of

---

[5] Note that the first 5 rows reporting deviations (not absolute deviations) would be the same for the pooled results as for the average of the sample estimates that are reported.

[6] A random selection from the columns of Table 7.1.A.1 was made and the third column chosen for these experiments.

[7] As explained in the last chapter, the SF results are normalized so that the most efficient firm registers at 100% efficient. In fact, this makes only trivial difference. The issued is discussed in the appendix to the chapter in which the comparisons with "un-normalized" SF estimation are reported.

the distribution[8]. These results can be seen in all the individual samples listed in Tables 7.1.A.1 and 7.1.A.2. Generally, SF performs somewhat better on these measures.

In comparing the average absolute deviations, we find that SF performs better than DEA in all sample estimations. The performance difference varies from .0078 to 0.0206 in the first part of experiment (Table 7.1.A.1) and from 0.0114 to 0.0291 in the second part of experiment (Table 7.1.A.2). On average, the absolute deviation for the SF method is about two percent and for the DEA is about three percent (the average of the samples in both tables). In the pooled estimation, where the effect of random noise is partially removed, we find that the performance difference is somewhat smaller with a factor of about 1.5 percent in terms of the average absolute deviation (Table 7.1.A.1).

The Spearman's rank correlation coefficients indicate the relationship between the true ranks of the observations and the ranks of the estimates. For SF estimation, the coefficients vary from 0.8994 to 0.9365 in the first part of experiment and from 0.8640 to 0.9354 in the second part. For DEA estimation, the Spearman's correlation coefficients are somewhat lower in each sample. However, in the pooled

---

[8]. One would expect this tendency, given that the random noise component will "push the estimated frontier out".

estimation (after some of the random noise is averaged out) we find the SF and DEA yield quite similar results.

The Chi-square statistic reports on the ability of the methods to extract a correct efficiency profile, and we find that in most cases, SF has a better performance than DEA though both survive the goodness-of-fit test at 90% confidence level.

Case B: Efficiency Profile Type B.

In this case, the distribution of the efficiency term **t** is truncated in such a way that only one tail is cut off. In comparison with the Case A, more information about the original distribution is thus preserved. Thus, the SF model should have a better chance to attain more accurate estimation. This experiment set also contains two parts parallel to the Case A above: in the first part one drawing of efficiency index was merged with five sets of CRS production data; in the second part, five drawings of the efficiency index ware merged with one set of CRS production set. The results are reported in Table 7.1.B.1 and Table 7.1.B.2..

In examining the "individual index estimation", we do not find much evidence of systematic over- or under-shooting behaviour for DEA. On the other hand, it seems that SF has more chance to overshoot the targets. In the first part of the experiment, the average deviation of estimation is -0.0102 for the SF and is -0.0060 for the DEA. The difference factor is about 0.0042 or 0.4 percent (Table 7.1.B.1). In the second part of experiment, the figure is -0.0037 for the SF and 0.0041 for the DEA. The difference factor is about 0.0004 or 0.04 percent. However, since both over- and under-shooting occur, the signed "average deviation" is more volatile, and hence less indicative, than the unsigned "average absolute deviation".

In the first part of the experiment, the average absolute deviation is 0.0187 for the SF model and 0.0217 for the DEA model (Table 7.1.B.1). The difference factor here is about 0.3 percent. In the second part of the experiment, the figures are 0.0216 and 0.0235 (Table 7.1.B.2) respectively with a difference factor of 0.31 percent. In the pooled results, with some random noise being removed, we find that the figure is 0.0123 for the SF model and is 0.0113 for the DEA (the pooled average absolute deviation in Table 7.1.B.1). Thus for both methods, the "pooled" absolute estimation error is about one percent.

In this second profile (type B), the Spearman's rank correlation coefficients indicate that the rank of the efficiency index can be reconstructed with higher precision by both models (than for profile A). In the sample estimated, the coefficients for the SF results vary from 0.9876 to 0.9919 in the first part of experiment and from 0.9834 to 0.9914 in the second part of experiment. On average, the coefficients for the SF estimation are slightly higher than those for the DEA method. The difference factor between the two coefficients varies from 0.0021 (sample 2) to 0.0084 (sample 4) in the first part of experiment (Table 7.1.B.1) and from 0.0013 (sample 3) to 0.0072 (sample 4) in the second part of the experiment (Table 7.1.B.2). For the pooled estimations, the coefficient is 0.9974 for the SF model and 0.9971 for the DEA estimation. The difference is narrowed down to 0.0003.

In measuring the ability of efficiency profile recovery, we note that the Chi-square values for both the SF and DEA models are well below the critical value for the goodness-of-fit test at 90% confidence level (34.3816). Generally, DEA has better performance than SF. However, for the pooled estimation, the Chi-square value is 10.2579 for the SF model and 10.6173 for the DEA estimation.

This second profile (type B) of the experiment set I will be treated as a norm for the rest of the experiments. Moreover, since we have found that the two parts of the experiment yield much the same information (Table 7.1.A.1 v.s. 7.1.A.2 and Table 7.1.B.1 v.s. 7.1.B.2), our future experiments will concentrate on one type of experiment, namely, the samples are formed by 1 drawing of an efficiency index and 5 drawings of production data and random noise. Finally, only the average of the sample estimation (corresponding to the last two columns in Tables 7.1.A.1 and 7.1.B.1) will be reported for the sake of simplicity though we will report the Spearman's r and the Chi-square statistics for each of the samples.

Case C. Efficiency Profile Type C.

Unlike the previous two cases, the efficiency profile in Case C is pre-selected in such a way that it features a larger standard deviation for the efficiency index exp(-t).[9] The distribution is still some what like a truncated normal and can be treated as "approximately truncated normal". Since the selection procedure blurs the connection between the sample and the population, the sample rather than population statistics are reported in the title of Table 7.1.C. We note that the consistency properties of the ML procedure no longer hold when the efficiency profile is not a truncated normal. Nevertheless, it is worth noting whether the efficiency estimates from the SF degenerate if the profile is not truly a truncated normal.

In this case, we find that while both the SF and the DEA model have a higher chance to underestimate the efficiency index. On average, the SF estimation has an average deviation of -0.0147 from the target while DEA estimation has one of -

---

[9] It should be noted that both mean and standard deviation of exp(-t) are the functions of the mean and standard deviation of t. See Chapter 6.

0.0123 (Table 7.1.C). The average absolute deviation is 0.0303 for the SF estimation and is 0.0354 for the DEA. The difference factor here is 0.0051. For the pooled estimation, the average absolute deviations are 0.0198 and 0.0219 for the SF and DEA estimation respectively and the difference factor is 0.0021.

In Case C, the rank order of the original index is quite well reconstructed by both methods. However, there is evidence suggesting that unusual behaviour of a particular data set may exert considerable impact on both SF and DEA estimation. For example, in sample 3, the Spearman's r is 0.8720 for SF estimation and 0.8370 for DEA estimation. Both figures are substantially lower than the other samples estimated.[10] For pooled estimation, the Spearman's r is 0.9918 for the SF estimation and 0.9874 for the DEA estimation.

Measuring the efficiency profile recovery, we find in general the SF performs better. In four out of the five samples, the SF yields a better Chi-square value. The difference between the pairs of Chi-square statistics varies from 1.684 to 9.5096. However, for the pooled estimation, the Chi-square value is 17.4933 for the SF estimation and 10.3352

---

[10] Since the DEA and the SF estimation are carried out independently, it seems obvious that the source of variation here is in the data set.

for the DEA estimation, an insignificant difference.

In comparison to the previous cases, the quality of the SF estimation is relatively less accurate. One possible reason is that the larger variance of the distribution makes the frequency distribution thicker and somewhat featureless. As a result, the ML procedure may be less able to successfully pinpoint to the maximum which is not very distinguishable from its neighbours.

Case D. Efficiency Profile Type D.

The efficiency profile of Case D features an even further deviation from a truncated normal. In fact it is an asymmetrical distribution for the efficiency index exp(-t). It skews towards the left with the mean located at 0.63 and a sample standard deviation of 0.16 (see Figure 7.1). The efficiency index scatters widely over a range from about 0.3 though 1.0. The results of the experiment are reported in Table 7.1.D.

In the first group of statistics, we find that the average deviation of the SF estimation is larger than for the DEA (-0.0151 v.s. 0.0060). On the average, SF has more chances to undershoot the target while DEA is more likely to overshoot the target. The average absolute deviation is 0.0190 for the SF and 0.0233 for the DEA. In the pooled estimation (which eliminates some of the random noise), we find that the average absolute deviation is 0.0153 for the SF estimation and 0.0108 for the DEA. The difference factor is 0.0045.

Both approaches yield acceptable rank statistics. However, in each sample estimation, SF leads the competition by a small margin. For the pooled estimation, the spearman's r is 0.9983 for the SF model and is 0.9962 for the DEA model.

Compared to the previous cases, the fitting of the efficiency profile, as measured by Chi-square, for the SF model is somewhat less accurate in individual sample estimations. In three out of the five samples, the Chi-square values for the SF estimation exceed those of the DEA estimation. For the pooled estimation, the DEA estimation now performs better than the SF. The chi-square value is 20.4569 for the SF and is 7.4492 for the DEA. As one might expect, as the distribution of the efficiency term deviates further from the assumed truncated normal, SF has a relatively worse performance.

We draw the following conclusion from Experiment Set I:

(1) Under assorted efficiency profiles, the efficiency index can be estimated by either the SF or DEA model with reasonable accuracy. On average, the estimation errors are about 3 percent, and the SF model performs somewhat better. The difference in performance between the two methods

is small. We observe that the average difference between the two estimates is between one and two percent. After the effects of random noise are partially averaged out (for pooled estimation), the difference is even smaller. In Cases A and C, the individual index estimation by the SF method is not as accurate as in the other two cases. To account for these facts, we note that in Case A the ML procedure, in fact, attempts to infer the information about the entire distribution from a small tail. This could result in less accurate parameter estimation and hence in efficiency index estimation. For Case C, it should be noted that the large variance of the distribution may make the maximum more indistinguishable from its neighbourhood and hence may increase the difficulty of estimation.

(2). In all cases, the efficiency rank orders of observations can be reconstructed with considerable accuracy by both SF and DEA methods. In most cases, the rank correlation coefficients between the "true" and estimated indexes are higher than 0.9. However, if the efficiency indexes are more closely clustered as in Case A, the ranking of the estimated indexes may be less accurate. It is conceivable that in such cases a certain estimation error may result in larger disturbances in the efficiency ranking order.

(3) In all four cases, the efficiency profile can be fitted reasonably well by both methods. Though the SF model

has a better performance in general, we find occasionally that DEA yields better results. In spite of the fact that the Chi-square statistics for DEA are usually higher than those for SF, they are less volatile.

7.2 Experiment Set II: Specification Errors Regarding Production Technology.

The production structure defined in Chapter 2 allows that a production process has a property of either constant or decreasing returns to scale (DRS). Increasing returns to scale technology is ruled out due to its inconsistency with the convexity assumption. Thus, we assume in this set of experiments that production data are generated from either one of the two permissible technologies.

Pre-specification of the production technology is required for DEA. The specification errors emerge in the following two circumstances: **a)** a CRS technology is loosely specified as the Non-increasing Returns to Scale (NIRS) one, or **b)** a DRS technology is erroneously treated as a CRS one. Besides the two types of errors just referred to, the consequence of employing a VRS specification (which violate the basic axioms of production structure by containing a possible IRS portion) is also considered in our evaluation.

Implication of this experiment for the SF model should be noted in particular. Though pre-specification of production technology is not required, the SF method faces a harsher task to distinguish the technical inefficiency from the effect of returns to scale. In many cases, effect of decreasing returns to scale has similar impact to inefficiency on the output.

To explore the consequences of the technology mis-specification, we create two hypothetical data sets: one has a CRS property while the other has a DRS property. The drawing of the efficiency indexes in this experiment is identical to the one employed in Case B of the experiment set I.[11] The results of this experiment are reported in Table 7.2.A and Table 7.2.B.

Case A: CRS Production Technology.

In this case, the input and output data are generated as earlier from a CRS technology:

$$u = L^{0.4}K^{0.6}\ e^{v-t}$$

The samples are generated in such a way that one set of the efficiency index is merged with five sets of CRS production data and drawings of the random error. For each of the five samples, four estimation procedures (the SF model, CRS, NIRS and VRS versions of the DEA model) are applied.

[11] By retaining the results from case I.B, this choice provides both experiments with a comparable starting points.

Since the set of efficiency index is identical to the one employed in Case B of experiment set I, the first two columns in Table 7.2.A (the SF and CRS estimations) are the duplications of the corresponding columns in Table 7.1.B.1.

Columns 3 and 4 of the Table 7.1.B.1 are the test statistics for the DEA estimation as the NIRS or VRS technology is specified. To examine the individual index estimation, the column by column comparison shows that in comparison with the CRS model, the NIRS and VRS model have a tendency to "over-shoot the target". This result is consistent with the theoretical anticipation that both NIRS and VRS models may over-estimate efficiency index since both the NIRS and VRS frontier are nested below the CRS frontier. From the average absolute deviation measures we find that while the SF result (0.0187) is better than all the DEA models, the CRS model performs better than the NIRS and VRS models (0.0217 v.s. 0.0244 and 0.0427). However, for the pooled estimation, the NIRS model is slightly better than the CRS and the VRS model. It should be noted that while the difference between the CRS and the NIRS model are very small, the difference between them and the VRS model is quite significant.

The rank information exhibits a similar scenario. While the SF model keeps best track of the true rank order, the rank coefficient for the CRS is slightly higher than the one for the NIRS and the one for the VRS estimation. For the pooled estimation, we find that the correctly specified CRS model (Spearman's rank correlation coefficient is 0.9971) performs slightly better than the NIRS model (0.9966) and both of them perform much better than the VRS model (0.9782).

In comparing the Chi-square statistics across the samples, we find the CRS and the NIRS model yield very close results. However, the VRS fitting is a more distorted one. For the pooled estimation, the Chi-square statistic for the CRS estimation (10.2579) is better than for the NIRS model (12.9010) and for the VRS model (14.5595).

Case B: DRS Production Technology

The production function in this case has following specification:

$$u = L^{0.3} k^{0.4} e^{v-t}$$

As explained in Chapter 6, the proper DEA specification for the data set generated from this DRS structure is the NIRS model. CRS and VRS are now mis-specified models. The statistics of the estimation are listed in Table 7.2.B, which report the same information as in Table 7.1.A.

First we look at the individual index estimation. In comparison with the previous case, we find that the SF results are insignificantly changed. Though the average deviation for the SF model has a sign change (0.0005 in this case v.s. -0.0102 in previous case), the average absolute deviation (0.0180) is quite close to the previous one (0.0190). However, there is a change in the pooled average absolute deviation for the SF estimation (0.0071 in this case v.s. 0.0123 in previous case). The results indicate that the SF model can effectively distinguish the effect of technical inefficiency under alternative returns to scale structure.

In comparing the three DEA models, our observations are as follows: first, the CRS model has a tendency to underestimate the efficiency index when the structure is, in fact, DRS. This phenomenon is consistent with the theoretical anticipation because the CRS frontier envelopes the NIRS and VRS frontiers and hence efficiency for an observation will be underestimated if the CRS envelope is referenced. Further, we note that the correctly specified NIRS model ties with the CRS model and both of them perform better than the VRS model. In terms of the pooled estimation, the NIRS model performs better than the CRS and the VRS model.

Regarding the rank information, the column by column comparison shows that the ranking by the SF results continues to dominate the competition. Among the DEA models, the correctly specified NIRS model yields the best results (the rank correlation coefficient for the pooled estimation is

0.9953) though the CRS model, in spite of the mis-specification, still provides reasonable results. The VRS results, in contrast, are much weaker. These observations can be found in all sample estimations.

Regarding efficiency profile fitting, we find that the SF model still leads the competition with lower Chi-square values. However, the correctly specified NIRS model has a relatively worse profile fitting than the CRS model. Finally, the VRS estimation has the worst fitting of the efficiency profile (25.3810).

The experiment set II can be summarized as follows:

The correct specification of returns to scale technology is vitally important to DEA estimation. According to our experiments, the estimation result of a DEA model is sensitive to the technology specification. A correctly specified DEA model generally has a better performance. However, if a choice is between a CRS and a NIRS technology, the performance gap between the correctly and the incorrectly specified model is insignificant. Our experiments show that, in such a case, even a mis-specified model may yield a reasonable result. Nevertheless, if the VRS property were imposed on a data set which is generated from either the CRS or the NIRS world, the performance of the DEA could be impaired severely. The efficiency indexes would be over-estimated and the efficiency ranking and profile could be enormously distorted. This finding is consistent with the emperical observations by both Deprins et al. (1985) and Bjurek et al. (1990), that VRS specification yielded a set of efficiency estimates significantly different from ones obtained from other models or other versions of the DEA model. Thus discretion is strongly advised for employing a VRS model, unless a data set is assured to have the VRS property. In this regard, we believe that the SF model could have a role to assist the DEA model in correctly specifying an unknown technology.

Finally, the experiments indicate that the SF model can successfully distinguish the effect of inefficiency from the effect of returns to scale and performs better in both cases.

7.3 Experiment set III:
Varying the Random Noise Level

The standard deviation of the random variable **v** was set at 0.03 through all previous cases.[12] This noise level (standard deviation) can be considered as roughly a 3 percent disturbance in addition to the effect of technical

---

[12] This is the population parameter. The actual sample standard deviation may vary from sample to sample.

inefficiency. In experiment set III, we set the random noise to different levels in order to study the robustness of the estimation techniques in the pure stochastic disturbance.

We examine four cases here with the standard deviation being set at 0.01 in Case A, 0.03 in Case B, 0.06 in Case C and 0.09 in Case D. The experimental results are summarized in Table 7.3.A, Table 7.3.B, Table 7.3.C and Table 7.3.D. It should be noted that the efficiency profile employed in this experiment is identical to the one in Case B of the experiment set I. Therefore, Table 7.3.B duplicates the corresponding contents of Table 7.1.B.1.

In this experiment set, we discuss the four cases together for the sake of convenience. First we look at the individual efficiency estimation. As the noise level rises, we find that the estimation of the SF and the DEA model are affected in a somewhat different pattern. At the lowest noise level, Case A, DEA shows a tendency to over-estimate the efficiency index (Table 7.3.A). The over-estimation can presumably be explained by the inner envelope property of DEA's linear frontier.[13] However, the DEA frontier is pushed up by the random noise as its level rises, and as a result, under-estimation becomes a prevailing phenomenon (the second column in Table 7.3.B, 7.3.C and 7.3.D).

---

[13] This property surfaces only when the noise level is set at the lowest level.

On the other hand, the SF model may over- or under-estimate the efficiency indexes and the results seem to vary from case to case. From the average absolute deviation, we observe the followings; first, the SF in general has a smaller absolute deviation than the DEA in all four cases, and second, performance gap becomes smaller from Case A to B and then becomes larger in Case C and D. The differences between the two are; 0.0063 in the Case A, 0.0030 in the Case B, 0.0052 in Case C and 0.0134 in the Case D. For the pooled estimation the gaps are 0.0078 for the Case A, 0.0010 for the Case B, 0.0062 for the Case C and 0.0083 for the Case D. Thus, we consider in terms of individual efficiency index estimation, increased random noise seems to affect the estimation quality of both methods, though in a slightly different pattern.

The experiment shows that there is direct relationship between the random noise level and the ability of the rank order reconstruction. As the random noise level rises, the Spearman's rank correlation coefficients for both methods decrease. However, one may observe that the SF model leads the competition by a slight margin in each single case. Another point that should be noted here is that the rank order seems to be a quite stable measure. Even when the data set is severely contaminated by random noise, both SF and DEA methods could preserve a highly authentic efficiency ranking.

Regarding the ability of the methods to recover an efficiency profile, we find that though the SF model provides a better fit of the efficiency profile than the DEA at low noise level, its ability is more sensitive to the random noise disturbance. In particular, at the lowest noise level, SF provides a closer fit for all 5 samples but at the highest noise level, SF is better in only two out of five samples. However, for the pooled estimation, we find the SF still provide a better fitting of the give efficiency profile in most of the cases.

Experiment set III can be summarized as follows; the presence of increased random noise affects the estimation quality of both the SF model and the DEA model. Though with SF there is a built-in noise filtering mechanism which allows for noise, the superiority of the SF model to the DEA model is not a overwhelming one. We find that the performance of the SF model degenerates as the random noise level rises. Our experiments show that at least in some aspects, e.g., individual index estimation and profile estimation, the DEA model has acceptable performance even when the noise level becomes quite high. Moreover, at low noise levels, where one might expect that DEA models to perform best, the efficiency profile fitted by the DEA model is not as good as by the SF model, though it is a more stable one. Therefore, we conclude over various noise level, the DEA model remains a competitive alternative to the SF model.

7.4 Summary

Generally speaking, the efficiency estimation quality of the SF and the DEA models are quite competitive. However, our experiments show that SF estimation frequently leads the alternative by a small margin. To account for this fact, we note that DEA has a congenital disadvantage in handling random noise. Moreover, some credit has to be granted to the DEA model for the following unfavourable situation it faces; the production data are generated from a parametric production function for which the DEA model provides only an approximation at best. Taking these into account, the general performance difference between the two seems quite small. Thus under many circumstances, the DEA model and the SF model are competitive alternatives. Finally, we emphasise that the estimation cost, which confines our present scale of experiments, should be taken into consideration when a choice is made between alternative models. This is particularly true for the DEA model. For a given data set, a linear programming procedure has to be repeated for each observation and the size of the LP problem varies as the numbers of observations increase. As a result, the cost of estimation rises as the sample size increases. On the other hand, if numbers of observations are limited, the SF model may not yield satisfactory results. In this situation, DEA is a good alternative.

Table 7.1.A.1 Comparison of SF and DEA. Profile A
Experiment Set I: 1 efficiency index merged with 5 sets of CRS production data
Distribution Characteristics: mean[exp(-t)] = 0.904, std[exp(-t)] = 0.069, std[exp(-v)] = 0.03

	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5		Average of Sample Estimation	
Statistics	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA
Individual Index Estimation												
Average Deviation:	-0.0013	-0.0324	-0.0011	-0.0179	-0.0050	-0.0380	0.0009	-0.0160	-0.0073	-0.0238	-0.0028	-0.0256
Average Deviation for:												
25% most efficient obs.:	-0.0143	-0.0341	-0.0148	-0.0256	-0.0126	-0.0313	-0.0151	-0.0261	-0.0177	-0.0216	-0.0149	-0.0277
middle 50% of obs.:	0.0002	-0.0339	0.0015	-0.0167	-0.0055	-0.0430	0.0033	-0.0135	-0.0046	-0.0253	-0.0010	-0.0265
25% least efficient obs.:	0.0067	-0.0297	0.0059	-0.0145	0.0018	-0.0365	0.0102	-0.0129	-0.0043	-0.0252	0.0041	-0.0238
Average Absolute Deviation:	0.0173	0.0379	0.0182	0.0279	0.0193	0.0419	0.0223	0.0301	0.0216	0.0331	0.0197	0.0342
Pooled Average absolute Deviation:											0.0102	0.0263
Rank Information												
Spearman's r:	0.9365	0.8989	0.9219	0.8911	0.9113	0.9107	0.8994	0.8559	0.9058	0.8958	0.9772	0.9722
Efficiency Profile Information												
Chi-square (34.3816):	18.3706	28.4768	13.8487	18.2927	10.9353	36.7602	17.8044	12.2040	19.4453	19.2734	17.7480	28.0726

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation – see text for details.

Table 7.1.A.2 Comparison of SF and DEA. Profile A
Experiment Set I: 5 sets of efficiency indexes merged with 1 set of CRS production data
Distribution Characteristics: m[exp(-t)] = 0.904, std[exp(-t)] = 0.069, std[exp(-v)] = 0.03

	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5		Average of Sample	
Statistics	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA
Individual Index Estimation												
Average Deviation:	-0.0050	-0.0380	-0.0022	-0.0321	-0.0036	-0.0434	-0.0101	-0.0273	0.0027	-0.0257	-0.0036	-0.0333
Average Deviation for:												
25% most efficient obs.:	-0.0120	-0.0313	-0.0172	-0.0349	-0.0178	-0.0437	-0.0244	-0.0323	-0.0120	-0.0306	-0.0167	-0.0346
middle 50% of obs.:	-0.0055	-0.0430	0.0013	-0.0338	0.0001	-0.0430	-0.0082	-0.0270	0.0099	-0.0221	-0.0005	-0.0338
25% least efficient obs.:	0.0018	-0.0365	0.0055	-0.0264	0.0030	-0.0442	0.0002	-0.0233	0.0037	-0.0203	0.0028	-0.0301
Average Absolute Deviation:	0.0193	0.0419	0.0198	0.0385	0.0204	0.0495	0.0225	0.0339	0.0195	0.0329	0.0203	0.0393
Rank Information												
Spearman's r:	0.9113	0.9107	0.9354	0.9206	0.9219	0.8724	0.8640	0.8523	0.8959	0.8908	n.a.	n.a.
Efficiency Profile Information												
Chi-square (34.3816) :	10.9353	36.7602	6.8956	17.6524	11.6649	19.4300	19.2826	20.7675	15.8127	30.7451	n.a.	n.a.

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.

Table 7.1.B.1 Comparison of SF and DEA. Profile B

Experiment Set I: 1 efficiency index merged with 5 sets of CRS production data

Distribution Characteristics: mean[exp(-t)] = 0.718, std[exp(-t)] = 0.138, std[exp(-v)] = 0.03

	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5		Average of Sample Estimations	
Statistics	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA
Individual Index Estimation												
Average Deviation:	-0.0136	-0.0011	-0.0088	-0.0123	-0.0082	0.0014	-0.0102	-0.0078	-0.0100	-0.0100	-0.0102	-0.0060
Average Deviation for:												
25% most efficient obs.:	-0.0142	0.0068	-0.0158	-0.0173	-0.0080	0.0073	-0.0165	-0.0104	-0.0161	-0.0124	-0.0141	-0.0052
middle 50% of obs.:	-0.0154	-0.0033	-0.0077	-0.0102	-0.0095	0.0005	-0.0095	-0.0091	-0.0126	-0.0141	-0.0109	-0.0072
25% least efficient obs.:	-0.0095	-0.0050	-0.0044	-0.0012	-0.0058	-0.0029	-0.0054	-0.0028	0.0011	0.0002	-0.0048	-0.0023
Average Absolute Deviation:	0.0213	0.0234	0.0186	0.0213	0.0171	0.0186	0.0175	0.0239	0.0188	0.0216	0.0187	0.0217
Pooled Average absolute Deviation:											0.0123	0.0113
Rank Information												
Spearman's r:	0.9876	0.9814	0.9898	0.9877	0.9914	0.9893	0.9919	0.9835	0.9895	0.9858	0.9974	0.9971
Efficiency Profile Information												
Chi-square (34.3816) :	13.7821	14.8895	14.1872	12.2504	17.0283	20.9222	20.9296	19.5052	18.4112	10.5218	10.2579	10.6173

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.

(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.

(3) The results in the box refer to pooled estimation – see text for details.

Table 7.1.B.2 Comparison of SF and DEA. Profile A

Experiment Set I: 5 sets of efficiency indexes merged with 1 set of CRS production data

Distribution Characteristics: mean[exp(-t)] = 0.718, std[exp(-t)] = 0.138, std[exp(-v)] = 0.03

	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5		Average of Sample	
Statistics	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA	SF	DEA
Individual Index Estimation												
Average Deviation:	-0.0082	0.0014	0.0261	0.0233	-0.0232	-0.0163	-0.0065	0.0015	-0.0066	0.0104	-0.0037	0.0041
Average Deviation for:												
25% most efficient obs.:	-0.0080	0.0073	0.0287	0.0309	-0.0294	-0.0177	-0.0126	-0.0022	-0.0063	0.0133	-0.0055	0.0063
middle 50% of obs.:	-0.0095	0.0005	0.0283	0.0235	-0.0227	-0.0170	-0.0043	0.0050	-0.0067	0.0101	-0.0030	0.0044
25% least efficient obs.:	-0.0058	-0.0029	0.0190	0.0151	-0.0179	-0.0136	-0.0050	-0.0019	-0.0069	0.0080	-0.0033	0.0009
Average Absolute Deviation:	0.0171	0.0186	0.0294	0.0283	0.0260	0.0230	0.0176	0.0250	0.0178	0.0224	0.0216	0.0235
Rank Information												
Spearman's r:	0.9914	0.9893	0.9834	0.9812	0.9891	0.9878	0.9910	0.9825	0.9897	0.9835	n.a.	n.a.
Efficiency Profile Information												
Chi-square (34.3816) :	17.0283	20.9222	20.7066	15.3013	18.3426	15.4435	24.6568	18.3649	15.6644	13.6055	n.a.	n.a.

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.

(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.

Table 7.1.C Comparison of SF and DEA. Profile C
Experiment Set I: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.67, std[exp(-t)]=0.21, std[exp(-v)]=0.03

	Average of Sample Estimation	
Statistics	SF	DEA
Individual Index Estimation		
Average Deviation:	-0.0147	-0.0123
Average Deviation for:		
25% most efficient obs.:	-0.0268	-0.0257
middle 50% of obs.:	-0.0139	-0.0099
25% least efficient obs.:	-0.0043	-0.0039
Average Absolute Deviation:	0.0303	0.0354
Pooled Average absolute Deviation:	$\boxed{0.0198}$	$\boxed{0.0219}$
Rank Information		
Spearman's r:	$\boxed{0.9918}$	$\boxed{0.9874}$
Sample 1:	0.9906	0.9910
Sample 2:	0.9945	0.9853
Sample 3:	0.8720	0.8370
Sample 4:	0.9863	0.9846
Sample 5:	0.9919	0.9920
Efficiency Profile Information		
Chi-square (34.3816):	$\boxed{17.4933}$	$\boxed{10.3352}$
Sample 1:	7.5808	9.2648
Sample 2:	11.5616	16.1781
Sample 3:	25.0878	34.5974
Sample 4:	18.9703	22.5310
Sample 5:	10.5208	8.1314

Note: (1) The mean and standard deviation of exp(-t) in the title are sample statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.1.D Comparison of SF and DEA. Profile D
Experiment Set I: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)] = 0.63,std[exp(-t)]=0.16, std[exp(-v)]=0.03

	Average of Sample Estimations	
Statistics	SF	DEA
Individual Index Estimation		
Average Deviation:	-0.0151	0.0060
Average Deviation for:		
25% most efficient obs.:	-0.0263	0.0089
middle 50% of obs.:	-0.0146	0.0055
25% least efficient obs.:	-0.0050	0.0039
Average Absolute Deviation:	0.0190	0.0233
Pooled Average absolute Deviation:	0.0153	0.0108
Rank Information		
Spearman's r:	0.9983	0.9962
Sample 1:	0.9923	0.9876
Sample 2:	0.9938	0.9717
Sample 3:	0.9923	0.9805
Sample 4:	0.9931	0.9924
Sample 5:	0.9915	0.9863
Efficiency Profile Information		
Chi-square (34.3816) :	20.4569	7.4492
Sample 1:	22.2100	10.7209
Sample 2:	17.5524	26.5424
Sample 3:	22.6463	18.8199
Sample 4:	22.3261	16.8699
Sample 5:	16.1352	17.5861

Note: (1) The mean and standard deviation of exp(-t) in the title are sample statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.2.A Comparison of SF and DEA. Profile B
Experiment Set II: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)] = 0.03

	Average of Sample Estimations			
Statistics	SF	CRS	NIRS	VRS
Individual Index Estimation				
Average Deviation:	-0.0102	-0.0060	0.0024	0.0278
Average Deviation for:				
25% most efficient obs.:	-0.0141	-0.0052	0.0023	0.0222
middle 50% of obs.:	-0.0109	-0.0072	0.0020	0.0296
25% least efficient obs.:	-0.0048	-0.0023	0.0030	0.0296
Average Absolute Deviation:	0.0187	0.0217	0.0244	0.0427
Pooled Average absolute Deviation:	0.0123	0.0113	0.0101	0.0305
Rank Information				
Spearman's r:	0.9974	0.9971	0.9966	0.9782
Sample 1:	0.9876	0.9814	0.9817	0.9418
Sample 2:	0.9898	0.9877	0.9846	0.9415
Sample 3:	0.9914	0.9893	0.9897	0.7771
Sample 4:	0.9919	0.9835	0.9634	0.9258
Sample 5:	0.9895	0.9868	0.9851	0.9138
Efficiency Profile Information				
Chi-square (34.3816):	10.2579	10.6173	12.9010	14.5595
Sample 1:	13.7821	14.8895	16.3969	23.8742
Sample 2:	14.1872	12.2504	12.0547	20.3229
Sample 3:	17.0283	20.9222	23.6711	39.5977
Sample 4:	20.9296	19.5052	17.2063	23.5317
Sample 5:	18.4112	10.5218	10.6175	27.3737

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.2.B Comparison of SF and DEA. Profile B
Experiment Set II: 1 efficiency index merged with 5 sets of DRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)]=0.03

Statistics	Average of Sample Estimations			
	SF	CRS	NIRS	VRS
Individual Index Estimation				
Average Deviation:	0.0005	-0.0047	0.0191	0.0209
Average Deviation for:				
25% most efficient obs.:	-0.0031	-0.0060	0.0179	0.0273
middle 50% of obs.:	0.0009	-0.0065	0.0221	0.0431
25% least efficient obs.:	0.0030	-0.0003	0.0135	0.0377
Average Absolute Deviation:	0.0180	0.0316	0.0281	0.0452
Pooled Average absolute Deviation:	0.0071	0.0150	0.0020	0.0381
Rank Information				
Spearman's r:	0.9978	0.9928	0.9953	0.9813
Sample 1:	0.9900	0.9638	0.9720	0.8950
Sample 2:	0.9919	0.9827	0.9887	0.9021
Sample 3:	0.9903	0.9719	0.9849	0.9610
Sample 4:	0.9900	0.9774	0.9846	0.8307
Sample 5:	0.9933	0.9685	0.9831	0.9174
Efficiency Profile Information				
Chi-square (34.3816) :	4.7061	10.2658	13.8726	25.3810
Sample 1:	10.0916	14.3954	18.4063	26.9501
Sample 2:	10.0996	20.3365	22.4495	32.5398
Sample 3:	14.6941	20.7561	19.0540	16.7895
Sample 4:	17.4808	11.3026	13.5766	22.7945
Sample 5:	10.2490	17.6240	19.6852	29.1080

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.3.A Comparison of SF and DEA. Profile B
Experiment Set III: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)]=0.01

	Average of Sample Estimation	
Statistics	SF	DEA
Individual Index Estimation		
Average Deviation:	0.0013	0.0114
Average Deviation for:		
25% most efficient obs.:	0.0016	0.0143
middle 50% of obs.:	0.0010	0.0118
25% least efficient obs.:	0.0014	0.0074
Average Absolute Deviation:	0.0068	0.0131
Pooled Average absolute Deviation:	0.0032	0.0110
Rank Information		
Spearman's r:	0.9995	0.9989
Sample 1:	0.9986	0.9954
Sample 2:	0.9988	0.9967
Sample 3:	0.9984	0.9944
Sample 4:	0.9982	0.9927
Sample 5:	0.9982	0.9959
Efficiency Profile Information		
Chi-square (34.3816):	2.4460	9.7833
Sample 1:	6.7797	10.7108
Sample 2:	6.7437	22.0967
Sample 3:	8.6285	24.0877
Sample 4:	8.8202	16.5305
Sample 5:	9.7951	18.6336

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.3.B Comparison of SF and DEA. Profile B
Experiment Set III: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)]=0.03

Statistics	Average of Sample Estimation SF	DEA
Individual Index Estimation		
Average Deviation:	-0.0102	-0.0060
Average Deviation for:		
25% most efficient obs.:	-0.0141	-0.0052
middle 50% of obs.:	-0.0109	-0.0072
25% least efficient obs.:	-0.0048	-0.0023
Average Absolute Deviation:	0.0187	0.0217
Pooled Average absolute Deviation:	$\boxed{0.0123}$	$\boxed{0.0113}$
Rank Information		
Spearman's r:	$\boxed{0.9974}$	$\boxed{0.9971}$
Sample 1:	0.9876	0.9814
Sample 2:	0.9898	0.9877
Sample 3:	0.9914	0.9893
Sample 4:	0.9919	0.9835
Sample 5:	0.9895	0.9868
Efficiency Profile Information		
Chi-square (34.3816):	$\boxed{10.2579}$	$\boxed{10.6173}$
Sample 1:	13.7821	14.8895
Sample 2:	14.1872	12.2504
Sample 3:	17.0283	20.9222
Sample 4:	20.9296	19.5052
Sample 5:	18.4114	10.5218

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.3.C Comparison of SF and DEA. Profile B
Experiment Set III: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)]=0.06

	Average of Sample Estimations	
Statistics	SF	DEA
Individual Index Estimation		
Average Deviation:	0.0029	-0.0196
Average Deviation for:		
25% most efficient obs.:	-0.0170	-0.0294
middle 50% of obs.:	0.0080	-0.0192
25% least efficient obs.:	0.0123	-0.0078
Average Absolute Deviation:	0.0338	0.0386
Pooled Average absolute Deviation:	0.0174	0.0236
Rank Information		
Spearman's r:	0.9939	0.9971
Sample 1:	0.9651	0.9535
Sample 2:	0.9678	0.9587
Sample 3:	0.9705	0.9644
Sample 4:	0.9683	0.9621
Sample 5:	0.9801	0.9805
Efficiency Profile Information		
Chi-square (34.3816):	19.4886	13.1976
Sample 1:	18.9115	13.3464
Sample 2:	25.3040	22.1231
Sample 3:	28.9478	18.0481
Sample 4:	22.3677	22.3457
Sample 5:	22.0759	23.8683

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

Table 7.3.D Comparison of SF and DEA. Profile B
Experiment Set III: 1 efficiency index merged with 5 sets of CRS production data
Dist. Characteristics: mean[exp(-t)]=0.72, std[exp(-t)]=0.14, std[exp(-v)]=0.09

	Average of Sample Estimations	
Statistics	SF	DEA
Individual Index Estimation		
Average Deviation:	0.0006	-0.0266
Average Deviation for:		
25% most efficient obs.:	-0.0348	-0.0426
middle 50% of obs.:	0.0121	-0.0215
25% least efficient obs.:	0.0128	-0.0212
Average Absolute Deviation:	0.0481	0.0615
Pooled Average absolute Deviation:	0.0259	0.0342
Rank Information		
Spearman's r:	0.9857	0.9820
Sample 1:	0.9383	0.9259
Sample 2:	0.9226	0.8699
Sample 3:	0.9284	0.9239
Sample 4:	0.9268	0.9161
Sample 5:	0.9453	0.9328
Efficiency Profile Information		
Chi-square (34.3816):	16.8519	23.7588
Sample 1:	17.0293	23.7825
Sample 2:	17.1631	18.7901
Sample 3:	23.9638	22.8242
Sample 4:	28.6321	22.4022
Sample 5:	20.3020	15.9332

Note: (1) The mean and standard deviation of exp(-t) in the title are population statistics.
(2) The bracketed Chi-square value is the critical value for the upper 10% tail of the Chi-square distribution.
(3) The results in the box refer to pooled estimation -- see text for details.

CHAPTER 8. A SUMMARY

Edwin Newman once noted:[1]

> *Inputs are of course everywhere. It is my observation that for some reason there are far more inputs than there are outputs, which means a large number of puts are disappearing in the process. God knows where they'll turn up.*

Apparently, pursuing efficiency is a fundamental characteristic of production activity. Searching for better measurement of productive and technical efficiency, therefore, seems to be an everlasting assignment for economists. This concern is a primary motivation of the present dissertation.

This dissertation concentrates on the two leading approaches to efficiency measurement developed recently, namely, the Data Envelopment Analysis (DEA) and the Stochastic Frontier production model (SF). The study, extending from the theoretical background to empirical simulation experiments, can be summarized as follows:

1. We reviewed extensively the theoretical background of the both approaches. From the basic axioms, we note that the structure of production can be interpreted in such a way that

---

[1] E. Newman. A Civil Tongue, New York: Warner Books, 1975. Quoted from R. Kopp (1981), 477.

inefficient operation is permissible behaviour. Thus an efficiency structure can be established based on this framework. This structure, namely the Debreu-Farrell efficiency structure, is the corner-stone underpinning our present analysis.

2. Among various efficiency measurement procedures, the stochastic frontier production model (SF) and data envelopment analysis (DEA) draw most attentions due to their distinctive features. Our analysis focuses on the following aspects of the two models:

(i). For the DEA model, we attempt to relate it to modern production analysis and thus put it on a solid basis. The propositions in Chapter 4, which are based on Afriat (1972), are the results of this effort. In a Lagrangian approach, we explored the nature of the optimal solution in the non-parametric programming model and gave a formal interpretation for its relation with the parametric approach. Furthermore, we demonstrated relationships between alternative DEA models that have different returns to scale properties.

(ii). In the SF approach, based upon the truncated-normal/normal model proposed by Stevenson (1980), and Jondrow et. al (1982), we worked out the conditional estimation procedure for the firm-specific efficiency measurement. The procedure provides more flexibility in handling a wider spectrum of efficiency profiles.

3. To compare the effectiveness of the DEA and SF models in regards to empirical application, we carried out a series of simulation experiments to test the two models in various aspects. The findings of the experiments can be briefly summarized as:

(i). Both models perform reasonably well in capturing assorted efficiency profiles.

(ii). Returns to scale specification is important to the DEA approach. The correctly specified model has a better performance than the mis-specified one. Moreover, the models with CRS and NIRS specification yield more similar results while the VRS model gives a worse one, providing that the convexity axiom is assumed. Further, we find that the SF model can successfully separate the effects of technical inefficiency and the effects of returns to scale.

(iii). The magnitude of random noise affects estimation accuracy. Though the SF model is designed with a built-in noise filter, an increasing noise level does have an adverse effect on its efficiency measurement. In comparison with the SF model, the DEA model in fact has quite similar performance across different noise levels.

Our study and experimental results suggest some unexplored issues deserving of further attention. First, in our experiments, we assume a simple production structure and no specification error for the SF model. However, the mis-specification of the production structure is a persistent threat to parametric frontier models. Thus the behaviour of the SF model under a harsher situation needs to be examined. Second, the large sample properties of both methods are subject to further evaluation. Confined by estimation cost, our analysis has focused on the finite sample properties of both models. However, one may expect that the increase of sample size may affect the estimation quality. For the SF model, large samples may help to eliminates the computation errors. On the other hand, though DEA is not a statistical model, but a larger sample size can help DEA to yield a refined linear approximation of a given frontier, provided the frontier does exist. Third, applications of both DEA and SF to panel data are worthy of further exploration. Applying the SF model to panel data has been a centre of interests in recent period. However, applying DEA to panel data should also be a premising area. Finally, complementary usage of the SF and the DEA methods should be emphasised in particular. By employing their own unique merits, two methods can work complementarily to identify the structure of data and the underlying technology, and give an observed data set a proper interpretation.

Upon finishing this review, Edwin Newman's comment quoted earlier is once more recalled. Though revealing the efficiency of production has been an overriding pursuit for economists, many aspects of the subject remain unknown to us. It is hoped, however, that the present dissertation may serve as a stepping-stone to further the endeavour.

Appendix 1: The Gradient Vector for the SF Model With Truncated-Normal/Normal Distribution

Provided technical inefficiency component is distributed as truncated-normal and random noise has a normal distribution, the logarithmic likelihood function corresponding to the marginal density function of the composite error is :

$$L(\varepsilon) = \sum_{\varepsilon} \ln \frac{e^{-\frac{(\varepsilon+\mu)^2}{2(\sigma_t^2+\sigma_v^2)}}}{\sqrt{2\pi}\sqrt{\sigma_t^2+\sigma_v^2}} \left[1 - F^{*}\left(\frac{\varepsilon\sigma_t^2 - \mu\sigma_v^2}{\sigma_t\,\sigma_v{*}\sqrt{\sigma_t^2+\sigma_v^2}} \right) \right]$$

where all the parameters are defined as in the text.

If production function has form:

$$y = u(\,x : \theta)\ e^{\varepsilon}$$

(ignoring the summation sign) the gradient vector for the likelihood function is:

$$\frac{\partial}{\partial\,\theta}\ L(\varepsilon) = \left\{ \frac{\sqrt{2}\ \sigma_t\ e^{-A^2/2}}{\sigma_v\sqrt{\sigma_t^2+\sigma_t^2}\ \sqrt{\pi}\ [erf(A/\sqrt{2}) - 1]} - \frac{\varepsilon+\mu}{\sigma_t^2+\sigma_v^2} \right\} \frac{\partial\varepsilon}{\partial\theta}$$

$$\frac{\partial}{\partial \mu} L(\varepsilon) = -\frac{\sqrt{2}\left\{\sigma_v \left[erf(\frac{\mu}{\sigma_t}) + 1\right] e^{-A^2/2} + \sqrt{\sigma_t^2 + \sigma_v^2}\,[erf(A/\sqrt{2}) - 1]\, e^{-\mu/2\sigma_t^2}\right\}}{\sigma_t \sqrt{\pi}\sqrt{\sigma_t^2 + \sigma_v^2}\,[erf(A/\sqrt{2}) - 1]\left[erf(\frac{\mu}{\sqrt{2}\,\sigma_t}) + 1\right]} - \frac{\varepsilon + \mu}{\sigma_t^2 + \sigma_v^2}$$

$$\frac{\partial}{\partial \sigma_t} L(\varepsilon) = \left\{ [erf(A/\sqrt{2}) - 1]\left[erf(\frac{\mu}{\sqrt{2}\,\sigma_t}) + 1\right]\sqrt{\pi}\,\sigma_t^2\,(\sigma_t^2 + \sigma_v^2)^{3/2}\right\}^{-1}$$

$$\sqrt{2}\left\{\sigma_v\left[\varepsilon\,\sigma_t^2 + \mu\,(2\sigma_t^2 + \sigma_v^2)\right]\left[erf(\frac{\mu}{\sqrt{2}\,\sigma_t}) + 1\right] e^{-A^2/2}\right.$$

$$\left. -\frac{\sqrt{2}}{2}\sqrt{\sigma_t^2 + \sigma_v^2}\,[1 - erf(A/\sqrt{2})]\left\{\mu\sqrt{2}\,(\sigma_t^2 + \sigma_v^2)\, e^{-\frac{\mu^2}{2\sigma_t^2}} - \sigma_t^3\sqrt{\pi}\left[erf(\frac{\mu}{\sqrt{2}\sigma_t}) + 1\right]\right\}\right\}$$

$$+ \sigma_t \frac{(\varepsilon + \mu)^2}{(\sigma_t^2 + \sigma_v^2)^2}$$

$$\frac{\partial}{\partial \sigma_v} L(\varepsilon) = \sigma_v \frac{(\varepsilon + \mu)^2}{(\sigma_t^2 + \sigma_v^2)^2}$$

$$- \frac{\sqrt{2}\left\{\sigma_t\,[\varepsilon\,(\sigma_t^2 + 2\,\sigma_v^2) + \mu\,\sigma_v^2)\,]\,e^{-A^2/2} + \frac{1}{\sqrt{2}}\sigma_v^3\sqrt{\sigma_t^2 + \sigma_v^2}\sqrt{\pi}\,[erf(A/\sqrt{2}) - 1]\right\}}{\sigma_v^3\,[erf(A/\sqrt{2}) - 1]\sqrt{\pi}\,(\sigma_t^2 + \sigma_v^2)^{3/2}}$$

where **erf()** is the error function and

$$A = \frac{\varepsilon\,\sigma_t^2 - \mu\,\sigma_v^2}{\sigma_t\,\sigma_v\sqrt{\sigma_t^2 + \sigma_v^2}}$$

This vector can be employed to build a Hessian matrix to guide optimal searching directions in the *Davidon-Fletcher-Powell* multidimensional nonlinear optimization algorithm.

Appendix 2. Normalization of Efficiency Indexes

Assume **t** is the efficiency term with a density function **p()**:

$$t \sim p(t;\theta), \qquad t \geq 0$$

where θ is the parameter set of the distribution. We define

$$z = e^{-t}, \qquad t \geq 0$$

so the **z** is a normalized efficiency index with an effective range **(0,1]**. However, for a sample drawing of **t,** we may have

$$\min\,[t] > 0, \quad \rightarrow \quad \max\,[\,z\,] = e^{-\min[t]} < 1.0$$

then a normalization procedure

$$z^{*} = \frac{z}{\max[z]} = e^{-t + \min[t]}$$

is suggested to yield at least one 100% efficiency observation. Following Schmidt ans Sickles (1984), let $\mathbf{t}^{*}$ = **t - min[t],** then the density function for $\mathbf{t}^{*}$ has same properties as **p(t)** if sample size can be increased. In the context of our experiments, if a sample of **t** is drawn from a truncated normal distribution, then a ML procedure is equally applicable to **t** and its transformation $\mathbf{t}^*$.

Appendix 3: The FORTRAN Code for DEA Estimation (CRS Model)

```
      SUBROUTINE USER(LIN)
C    *************************************************************
C    A FORTRAN subroutine interfaces with LINDO to conduct data
C    envelopment analysis (DEA). The programm is designed
C    exclusively for present dissertation. The prototypes of
C    the called subroutines (DEFROW,APPCOL) can be found in LINDO's
C    manual.
C    Input data format:     (3x,4(1x,f8.5))
C    Times of iteration:    100
C    Output mode:           TERSE (optional)
C    (c) Copyright: Dading Li  June,1991.
C    *************************************************************
C
      CHARACTER*1 KNAME,NULL,MALPH
      DIMENSION IRO(80)
      DIMENSION MALPH(12),KNAME(8)
      DIMENSION E(100),OUT(100),X(2,100)
      DIMENSION VAL(80)
      LOGICAL TRUBLE
      INFILE=9
      CALL LUNGET(INFILE,1,0)
      IF(INFILE .LE. 0) RETURN
C    DATA READING
      DO 22 I=1,100
      READ(INFILE,20) E(I),OUT(I),X(1,I),X(2,I)
        OUT(I)=E(I)*OUT(I)
20      FORMAT(3X,4(1X,F8.3))
22    CONTINUE
      DATA MALPH/'0','1','2','3','4','5','6','7','8','9','S','L'/
      DATA NULL/' '/
      TRUBLE=.TRUE.
      DO 900 N=1,100
        WRITE(5,23) N
23      FORMAT(10X,'CURRENT COUNT IS:',I3)
C     INITIALIZATION
      CALL INIT
C     OBJ FUNCTION
      CALL DEFROW(-1,0.,IDROW,TRUBLE)
C     CONSTRAINTS 1
      DO 100 I=2,3
      CALL DEFROW(1,X(I-1,N),IDROW,TRUBLE)
100   CONTINUE
C       CONSTRAINTS 2
        CALL DEFROW(1,0.,IDROW,TRUBLE)
C
C     CONSTRAINTS 3
        DO 200 I=1,100
      CALL DEFROW(-1,0.,IDROW,TRUBLE)
200   CONTINUE
C        LAST CONSTRAINT
C
```

```
C     GENERATING L1 TO L100
      KNAME(1) = MALPH(12)
      KNAME(4) = NULL
      KNAME(5) = NULL
      KNAME(6) = NULL
      KNAME(7) = NULL
      KNAME(8) = NULL
C
      DO 400 I=1,100
      M1=INT(I/10)
      M2=I-M1*10
      KNAME(2)=MALPH(M1+1)
      KNAME(3)=MALPH(M2+1)
      VAL(1)=X(1,I)
      VAL(2)=X(2,I)
      VAL(3)=-OUT(I)
        VAL(4)=1.0
      IRO(1)=2
      IRO(2)=3
        IRO(3)=4
        IRO(4)=I+4
        NONZ=4
      CALL APPCOL(KNAME,NONZ,VAL,IRO,TRUBLE)
400   CONTINUE
C
C     GENERATING L0
C
      KNAME(1)=MALPH(12)
      KNAME(2)=MALPH(1)
      KNAME(3)=NULL
      KNAME(4)=NULL
      KNAME(5)=NULL
      KNAME(6)=NULL
      KNAME(7)=NULL
      KNAME(8)=NULL
      IRO(1)=1
      IRO(2)=4
      VAL(1)=1.0
      VAL(2)=OUT(N)
      NONZ=2
      CALL APPCOL(KNAME,NONZ,VAL,IRO,TRUBLE)
        CALL LOOK(1,104)
      CALL GO(LIMGO,ISTAT)
900   CONTINUE
      STOP
      END
```

Appendix 4: The FORTRAN Code for the Maximum Likelihood Estimation (Truncated-normal/Normal SF model)

```
        PROGRAM  trucn
c*****************************************************************
c       This Program uses the Davidon-Fletcher-Powell gradient direction
c       multidimentional optimazition algorithm to miniminize the
c       negative maximum likelihood function 'func'. The initial parameter
c       vector is p(1...n). The gradient vector g(1...n) is derived by
c       calulating derivative function subroutine 'dfunc'. The searching
c       uses line  minimazation and golden section search techniques.
c       No initial bracketing for current problem.
c       Estimated parameters, standard deviation and inversed information
c       matrix are provided in output file 'report.txt'. Conditional mean
c       vector is printed in another output file 'cmean' [format(4x,f8.5)]
c       Input data file format (3x,4(1x,f8.5)).
c                  (C)  copyright:  Dading  Li,  June,  1991.
c*****************************************************************
        REAL HES(5,5),FTOL,FRET
        REAL EF(100),Y,L,K,RS
        REAL P(5), CM(100),MAX
        DOUBLE PRECISION ET(100)
        COMMON /AAA/Y(100),L(100),K(100)
        INTEGER N,ITER
        OPEN(UNIT=1,FILE='DAT',STATUS='OLD' )
        OPEN(UNIT=2,FILE='CMEAN',STATUS='NEW')
        OPEN(UNIT=3,FILE='REPORT.TXT',STATUS='NEW')
        WRITE(*,10)
        DO 1 I=1,100
                READ(1,2) EF(I),Y(I),L(I),K(I)
                Y(I)=LOG(EF(I)*Y(I))
                L(I)=LOG(L(I))
                K(I)=LOG(K(I))
1       CONTINUE
2       FORMAT(3X,4(1X,F8.5))
        N= 5
        FTOL=0.0001
        DO 3, I=1,N
        WRITE(*,6) I
        READ(6,*) P(I)
3       CONTINUE

        CALL DFPMIN(P,N,FTOL,ITER,FRET,HES)
        IF(P(4) .LT. 0.0001) THEN
           WRITE(*,911)
           GOTO 920
        ENDIF
        CALL CMEAN(P,ET,RS)
        WRITE(*,*) RS
        MAX=0.0
        DO 4, I=1,100
         CM(I)= EXP(-ET(I))
         WRITE(2,5) CM(I)
4       CONTINUE
```

```
5       FORMAT(4X,F8.5)
6       FORMAT(4X,'ENTER INITIAL VALUE FOR COEFFICIENT', 2X,I2)


10      FORMAT(5X,'READING INPUT FILE...')
        WRITE(3,100)
        WRITE(3,200)
        WRITE(3,300)
        WRITE(3,310)
        WRITE(3,*)
        WRITE(3,320)
        WRITE(3,*)
        WRITE(3,400)
        DO 50,I=1,N
        WRITE(3,500) I, P(I), SQRT(ABS(HES(I,I)))
50      CONTINUE
        WRITE(3,*)
        WRITE(3,590)
        WRITE(3,600) RS
        WRITE(3,*)
        WRITE(3,700)
        WRITE(3,*)
        DO 60,I=1,N
        WRITE(3,800) HES(I,1),HES(I,2),HES(I,3),HES(I,4),HES(I,5)
60      CONTINUE
        WRITE(3,*)
        WRITE(3,100)
100     FORMAT(10X,'**************************************************
     1************')
200     FORMAT(20X,'MAXIMUM LIKELIHOOD ESTIMATION REPORT')
300     FORMAT(20X,'     VARIABLE METRIC METHOD')
310     FORMAT(20X,'(DAVIDON-FLETCHER-POWELL ALGORITHM)')
320     FORMAT(10X,'MODEL SPECI.:                        AUTHOR: DADING LI')
400     FORMAT(35X,'ESTIMATION          ASYMPTOTIC STD')
500     FORMAT(10X,'COEFFICIENT',I2,':',10X,F8.5,12X,F8.5)
590     FORMAT(10X,'NUMBER OF OBSERVATIONS:             100')
600     FORMAT(10X,'SUM OF SQUARED RESIDUALS :',3X,F10.5)
700     FORMAT(10X,'INVERSED INFORMATION MATRIX (CRAMER-RAO MVB) :')
800     FORMAT(20X,5(2X,F8.5))
        WRITE(*,900)
        WRITE(*,910)
900     FORMAT(5X,'THE JOB IS DONE. PLEASE RENAME OUTPUT FILES ')
910     FORMAT(5X,'REPORT.TXT AND CMEAN TO PROTECT THE CONTENTS')
911     FORMAT(5X,'WARNNING: DATA SET IS NOT CONSISTENT WITH SF MODEL')
920     CLOSE(1)
        CLOSE(2)
        CLOSE(3)
        STOP
        END
C
        SUBROUTINE DFPMIN(P,N,FTOL,ITER,FRET,HESSIN)
        REAL FP,FAE,FAD,FAC,FRET
        REAL P(5),XI(5),FTOL
        REAL G(5),DG(5),HDG(5),HESSIN(5,5)
        INTEGER I,J,ITS
        ITMAX=40
        EPS=0.00000001
        CALL FUNC(P,FRET)
        CALL DFUNC(P,G)
        FP=FRET
        WRITE(*,1)
```

```
1       FORMAT(5X,'INITIAL PARAMETER VALUE :')
        WRITE(*,3) P(1),P(2),P(3),P(4),P(5)
        WRITE(*,2)
2       FORMAT(5X,'INITIAL GRADIENT VECTOR EVALUATION :')
        WRITE(*,3) G(1),G(2),G(3),G(4),G(5)
3       FORMAT(5X,5(2X,F8.5))
        DO 10,I=1,N
                DO 5,J=1,N
                        HESSIN(I,J)=0.0
                        IF(I .EQ. J) HESSIN(I,J)=1.0
5               CONTINUE
                XI(I) = -G(I)
10      CONTINUE
        ITS=0.0
20      ITS=ITS+1
        ITER=ITS
                WRITE(*,21) ITS,FRET
21          FORMAT(10X,'DPF ITERATION',I3,5X,'FUNCTION VALUE',F10.5)
                CALL LINMIN(P,XI,N,FRET)
           IF(HESSIN(4,4) .LT. 0.0) THEN
           HESSIN(4,4) = 0.0
           GOTO 25
           ELSEIF(HESSIN(5,5) .LT. 0.0) THEN
           HESSIN(5,5) =0.0
           GOTO 25
           ENDIF
           IF(2.0*ABS(FRET-FP) .LE. FTOL*(ABS(FRET)+ABS(FP)+EPS)) THEN
                RETURN
                ENDIF
25              FP=FRET
                DO 30,I=1,N
                        DG(I)=G(I)
30              CONTINUE
                CALL FUNC(P,FRET)
                CALL DFUNC(P,G)
                DO 40,I=1,N
                        DG(I)=G(I)-DG(I)
40              CONTINUE
                DO 50 I=1,N
                        HDG(I)=0.0
                        DO 45,J=1,N
                                HDG(I)=HDG(I)+HESSIN(I,J)*DG(J)
45                      CONTINUE
50              CONTINUE
                FAC=0.0
                FAE=0.0
                DO 60,I=1,N
                        FAC = FAC+DG(I)*XI(I)
                        FAE = FAE +DG(I)*HDG(I)
60              CONTINUE
                FAC=1.0/FAC
                FAD=1.0/FAE
                DO 70,I=1,N
                        DG(I)=FAC*XI(I)-FAD*HDG(I)
70              CONTINUE
                DO 80,I=1,N
                        DO 75,J=1,N
                                HESSIN(I,J)=HESSIN(I,J)+FAC*XI(I)*XI(J)
     1                                -FAD*HDG(I)*HDG(J)+FAE*DG(I)*DG(J)
75                      CONTINUE
80              CONTINUE
```

```
                DO 90 ,I=1,N
                        XI(I)=0.0
                        DO 85 J=1,N
                                XI(I) =XI(I)-HESSIN(I,J)*G(J)
85                      CONTINUE
90              CONTINUE
                IF( ITS .GT. ITMAX) THEN
                GOTO 100
                ENDIF
        GOTO 20
100     WRITE(*,101)
101     FORMAT( 2X,'TOO MANY ITERATION')
        RETURN
        END
C
        SUBROUTINE LINMIN(P,XI,N,FRET)
        REAL P(5),XI(5)
        REAL AX,XX,XMIN,CX
        REAL FRET,FIDIM,TOL
        INTEGER N
        COMMON PCOM(5),XICOM(5),NCOM
        WRITE(*,2)
2       FORMAT(10X,'LINMIN')
        TOL=0.001
        NCOM=N
        AX=-1.0
        XX=0.0
        CX=1.0
        DO 10, J=1,N
                PCOM(J) = P(J)
                XICOM(J)=XI(J)
10      CONTINUE
        CALL GOLDEN(AX,XX,CX,FIDIM,TOL,XMIN)
        DO 20,J=1,N
                XI(J) =XI(J)*XMIN
                P(J) =P(J) + XI(J)
20      CONTINUE
        FRET=FIDIM
        RETURN
        END
C
        REAL FUNCTION FIDIM(X)
        REAL X,FR
        REAL XT(5)
        COMMON PCOM(5),XICOM(5),NCOM
        DO 30, J=1,NCOM
                XT(J)=PCOM(J)+X*XICOM(J)
30      CONTINUE
        CALL FUNC(XT,FR)
        FIDIM=FR
        RETURN
        END
C
        SUBROUTINE GOLDEN(AX,BX,CX,F,TOL,XMIN)
        REAL AX,BX,CX,F,TOL,XMIN
        REAL COEF,FIDIM
        REAL F0,F1,F2,F3,X0,X1,X2,X3
        REAL R,C
        WRITE(*,2)
2       FORMAT(2X,'GOLDEN')
        R = 0.61803399
```

```
      C = 1.0-R
              X0 = AX
              X3 = CX
              IF(ABS(CX-BX) .GT. ABS(BX-AX)) THEN
                      X1=BX
                      X2=BX+C*(CX-BX)
              ELSE
                      X2=BX
                      X1=BX-C*(CX-AX)
              ENDIF
              COEF=X1
              F1=FIDIM(COEF)
              COEF=X2
              F2=FIDIM(COEF)
10            IF  (ABS(X3-X0) .LT. TOL*(ABS(X1)+ABS(X2))) THEN
                      GOTO 20
              ELSEIF (ABS(X3-X0) .LT. 0.000001) THEN
                      GOTO 20
              ELSE
                      IF(F2 .LT. F1) THEN
                              X0=X1
                              X1=X2
                              X2=R*X1+C*X3
                              F0=F1
                              F1=F2
                              COEF=X2
                              F2=FIDIM(COEF)
                      ELSE
                              X3=X2
                              X2=X1
                              X1=R*X2+C*X0
                              F3=F2
                              F2=F1
                              COEF=X1
                              F1=FIDIM(COEF)
                      ENDIF
              GOTO 10
              ENDIF
20            IF  (F1 .LT. F2) THEN
                      XMIN=X1
                      F=F1
              ELSE
                      XMIN=X2
                      F=F2
              ENDIF
      RETURN
      END
C
      SUBROUTINE FUNC(COEF,CF)
      REAL PI,SQ2 ,S
      REAL F,CF,E,Y,L,K,ERF
      REAL COEF(5), ALPHA,BETA,MU,SV,ST
      COMMON /AAA/Y(100),L(100),K(100)
      INTEGER N
      N=100
      PI=3.14159
      ALPHA=COEF(1)
      BETA=COEF(2)
      MU=COEF(3)
      ST=COEF(4)
      SV=COEF(5)
```

```
      SQ2=SQRT(2.0)
      S=ST**2+SV**2
      CF=0.0
      E=0.0
      IF(ST .EQ. 0.0) ST=0.00001
      IF(SV .EQ. 0.0) SV=0.00001
      D1=SQRT(2*PI*S)
      D2=1+ERF(MU/(ST*SQ2))
      D3=ST*SV*SQRT(S)*SQ2
      IF(D1 .EQ. 0.0) D1=0.000001
      IF(D2 .EQ. 0.0) D2=0.000001
      IF(D3 .EQ. 0.0) D3=0.000001
      DO 10,I=1,N
              E=Y(I)-ALPHA*L(I)-BETA*K(I)
              F=(1.0/(D1*D2))*EXP(-((E+MU)**2)/(2*S))
     $*(1.0-ERF((E*ST**2-MU*SV**2)/D3))
      IF(F .LE. 0.0) F=0.00001
      CF=CF-LOG(F)
10    CONTINUE
      IF(ST .LT. 0.0) CF=100000.0
      IF(SV .LT. 0.0) CF=100000.0
      RETURN
      END
C
      REAL FUNCTION ERF(X)
      DOUBLE PRECISION ANS

      REAL T,Z,X
      Z=ABS(X)
      T=1.0/(1+0.5*Z)
      ANS=(-Z*Z-1.26551223+T*(1.00002368+T*(0.37409196+T*
     $(0.09678418+T*(-0.18628806+T*(0.27886807+T*(-1.135204+T*
     $(1.48851587+T*(-0.82215223+T*0.17087277)))))))))
      IF(ANS .LT. -500.0) ANS=-500.0
      ANS=T*EXP(ANS)
      IF( X .GE. 0.0) THEN
               ERF=1.0-ANS
      ELSE
              ERF=1.0-(2.0-ANS)
      ENDIF
      RETURN
      END
C
      SUBROUTINE DFUNC(COEF,G)
      REAL G(5),Y,L,K
      REAL E,ALPHA,BETA,MU,G1,G3,G4,G5
      REAL CG1,CG2,CG3,CG4,CG5
      REAL SV,COEF(5),ERF,B,A
      COMMON /AAA/Y(100),L(100),K(100)
      INTEGER N
      N=100
      SQ2=SQRT(2.0)
      PI=3.14159
      ALPHA=COEF(1)
      BETA=COEF(2)
      MU=COEF(3)
      ST=COEF(4)
      SV=COEF(5)
      WRITE(*,*) ALPHA,BETA,MU,ST,SV
      G1=0
      G2=0
```

```
      G3=0
      G4=0
      G5=0
      CG1=0
      CG2=0
      CG3=0
      CG4=0
      CG5=0
      S=ST**2+SV**2
      B=MU/(SQ2*ST)
      ERB=ERF(B)
      IF(ERB .EQ. -1.0) ERB=-.999999
      DO 10,I=1,100
              E=Y(I)-ALPHA*L(I)-BETA*K(I)
              A=(E*ST**2-MU*SV**2)/(SV*ST*SQ2*SQRT(S))
              ERA=ERF(A)
              IF(ERA .EQ. 1.0) ERA=0.999999
          G1=SQ2*(ST**2)*EXP(-A**2)/(SV*(ERA-1)*SQRT(PI*S))-(E+MU)/S
              CG1=CG1+G1*(-L(I))
              CG2=CG2+G1*(-K(I))
              G3=-SQ2*(SV*(ERB+1.0)*EXP(-A**2)+SQRT(S)*(ERA-1.0)*
     $EXP(-B**2))/(ST*(ERA-1.0)*(ERB+1.0)*SQRT(PI*S))-(E+MU)/S
              CG3=CG3+G3
              G4=SQ2*(SV*(E*ST**2+MU*(2*ST**2+SV**2))*(1.0+ERB)*
     $EXP(-A**2)-(1.0/SQ2)*(SQRT(S)*(1-ERA)*(MU*SQ2*S*EXP(-B**2)-ST**3
     $*SQRT(PI)*(1+ERB)))  /  ((ST**2)*(ERA-1)*(ERB+1)*
     $SQRT(PI*(S**3))) +  ST*((E+MU)**2)/(S**2)
              CG4=CG4+G4
              G5=SV*((E+MU)**2)/(S**2) - SQ2*(ST*(E*(ST**2+2*SV**2)+
     $MU*SV**2)*EXP(-A**2) + (1.0/SQ2)*(SV**3)*SQRT(S*PI)*(ERA-1)) /
     $((SV**2)*(ERA-1)*SQRT(PI*(S**3)))
              CG5=CG5+G5
10    CONTINUE
      G(1)=-CG1
      G(2)=-CG2
      G(3)=-CG3
      G(4)=-CG4
      G(5)=-CG5
      RETURN
      END
C
      SUBROUTINE CMEAN(COEF,ET,RS)
      REAL Y,L,K
      REAL COEF(5),ALPHA,BETA,MU,ST,SV
      REAL E,ERF,SQ2,RS
      DOUBLE PRECISION DERF,ET(100),A
      COMMON /AAA/Y(100),L(100),K(100)
      RS=0.0
      PI=3.14159
      ALPHA=COEF(1)
      BETA=COEF(2)
      MU=COEF(3)
      ST=COEF(4)
      SV=COEF(5)
      S=SV**2+ST**2
      SQ2=SQRT(2.0)
      SQPI=SQRT(PI)
      SQS=SQRT(S)
      DO 10,I=1,100
              E=Y(I)-ALPHA*L(I)-BETA*K(I)
              A=(E*(ST**2)-MU*(SV**2))/(ST*SV*SQS)
```

```
            ET(I)=(SQ2*ST*SV*EXP(-0.5*A**2))/((1-DERF(A/SQ2))*SQPI*SQS)
     $ -(E*ST**2-MU*SV**2)/S
                R=(E+ET(I))**2
                RS=RS+R
10        CONTINUE

          RETURN
          END

          DOUBLE PRECISION FUNCTION DERF(X)
          REAL ANS,T,Z
          DOUBLE PRECISION X
          Z=ABS(X)
          T=1.0/(1.0+0.5*Z)
          ANS=(-Z*Z-1.26551223+T*(1.00002368+T*(0.37409196+T*
     $(0.009678418+T*(0.18628806+T*(0.27886807+T*(-1.135204+T*
     $(1.48851587+T*(-0.82215223+T*0.17087277))))))))))
          ANS=T*EXP(ANS)
          IF(X .GE. 0.0) THEN
             DERF=1.0-ANS
          ELSE
            DERF=1.0-(2.0-ANS)
          ENDIF
          RETURN
          END
```

BIBLIOGRAPHY

Afriat, S. N. "Efficiency Estimation of Production Functions", International Economic Review, 13 (1972),568-598.

Aigner, D.J. and S.F. Chu. "On Estimating the Industry Production Function", American Economic Review, 58 (4) (1968), 826-839.

Aigner, D.J., C.A.K. Lovell and P.J. Schmidt. "Formulation and Estimation of Stochastic Frontier Production Function Models", Journal of Econometrics, 6 (1977), 21-37.

Banker, R., A. Charnes and W.W. Cooper. "Models for the Estimation of Technical and Scale Inefficiencies in Data Envelopment Analysis", Management Science, 30 (9) (1984), 1078-1092.

Banker, R., A. Charnes, W.W. Cooper and A. Maindiratta. "A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology", A. Dogramaci and R. Fare (ed.). Applications of Modern Production Theory, Efficiency and Productivity, Boston: Kluwer Academic Publishers, 1988, Chapter 2.

Banker, R., A. Charnes, W.W. Cooper and A. Schinnar. "A Bi-Extremal for Frontier Estimation and Efficiency Evaluation", Management Science, 27 (12) (1981), 1370-1382.

Banker, R.D., R.F. Conrad and R.P. Strauss. "A Comparative Application of DEA and Translog Methods: An Illustrative Study of Hospital Production", Management Science, 32 (1) (1986), 30-44.

Banker, R. and A. Maindiratta. "Piecewise Loglinear Estimation of Efficient Production Surfaces", Management Science, 32 (1) (1986), 126-135.

Bauer, P.W. "Recent Developments in the Econometric Estimation of Frontiers", Journal of Econometrics, 46 (1990), 39-56.

Bjurek, H., L. Hjalmarsson and F.R. Forsund. "Deterministic Parametric and Nonparametric Estimation of Efficiency in Service Production - A Comparison", Journal of Econometrics, 46 (1990), 213-227.

Bury, K.V. Statistical Models in Applied Science. New York: John Wiley & Sons, 1975.

Charnes, A., W.W. Cooper, and E. Rhodes. "Measuring the Efficiency of Decision-Making Units", European Journal of Operation Research, 2(6) (1978), 429-444.

Charnes, A., W.W. Cooper and E. Rhodes. "Short Communication: Measuring the Efficiency of Decision Making Units", European Journal of Operation Research, (1979), 339.

Charnes, A., W.W. Cooper and E. Rhodes. "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions", Journal of Econometrics, 30 (1985), 91-107.

Charnes, A., W.W. Cooper, Z.M. Huang and D.B. Sun. "Polyhedral Cone-Ratio DEA Models with An Illustrative Application to Large Commercial Banks", Journal of Econometrics, 46 (1990), 73-91.

Debreu, G. "The Coefficient of Resource Utilization", Econometrica, 19 (3) (1953), 273-292.

Deprins, D., L. Simar and H. Tulkens. "Measuring Labour-Efficiency in Post Offices", in M.Marchand, P. Pestieau and H. Tulkens (ed.). The Performance of Public Enterprises: Concept and Measurement. Amsterdam: North -Holland, 1988.

Drake, A.W. Fundamentals of Applied Probability Theory. New York: McGraw-Hill Book Company, 1967.

Fare, R. Fundamentals of Production Theory. Berlin: Springer-Verlag, 1988.

Fare, R., S. Grosskopt and C.A. Lovell. The Measurement of Efficiency of Production. Boston: Kluwer Academic Publisher, 1985.

Farrell, M.J. "The Measurement of Productive Efficiency". Journal of the Royal Statistical Society, Vol. 120, Series A (General), Part 3 (1957), 253-290.

Farrell, M.J. and L. Fieldhouse. "Estimating Efficient Production Functions Under Increasing Return to Scale", Journal of the Royal Statistical Society, Series A. (1962), 252-267.

Ferrier, G.D. and C.A.K. Lovell. "Measuring Efficiency in Banking - Econometric and Linear Programming Evidence", Journal of Econometrics, 46 (1990), 229-254.

Forsund, F.R. and L. Hjalmarsson. "On the Measurement of Productive Efficiency", Swedish Journal of Economics (1974), 141-153.

Forsund, F.R. and E.S. Jansen. "On Estimating Average and Best Practise Homothetic Production Functions via Cost Functions", International Economic Review, 18 (1977), 463 -476.

Forsund, F.R., C.A.K. Lovell and P. Schmidt. "A Survey of Frontier Production Functions and of Their Relationship to Efficiency estimation", Journal of Econometrics, 13 (1980),1-25.

Greene, W.H. "Maximum Likelihood Estimation of Econometric Frontier Functions", Journal of Econometrics, 13 (1980), 27-56.

Greene, W.H. "A Gamma-Distributed Stochastic Frontier Model", Journal of Econometrics, 46 (1990), 141-163.

Johansen, L. Production Function. Amsterdam: North-Holland, 1972.

Jondrow, J., C.A.K. Lovell, I.S. Masterov, and P. Schmidt. "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", Journal of Econometrics, 19, (1982), 233-238.

Klein, L.R. A Textbook of Econometrics, Evanston: Row, Peterson and Co., 1953.

Koopmans, T.C. (ed). Activity Analysis of Production and Allocation. Cowles Commission for Research in Economics, Monograph No.13. New York: Wiley & Sons, 1951.

Kopp, R.J. "The Measurement of Productive Efficiency: A Reconsideration", Quarterly Journal of Economics, (1981), 477-503.

Kopp, R.J. and J. Mullahy. "Moment-Based Estimation and Testing of Stochastic Frontier Models", Journal of Econometrics, 46 (1990), 163-185.

Kumbhakar, S.C. "Production Frontiers and Panel Data: An Application to U.S. Class 1 Railroads", Journal of Business and Economic Statistics, 5 (1987), 249-255.

Kumbhakar, S.C. "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency", Journal of Econometrics, 46 (1990), 201-211.

Kumbhakar, S.C. and T. Summa "Technical Efficiency of Finnish Brewing Plants: A Production Frontier Approach", Scandinavian Journal of Economics, (1989), 149-160.

Lau, L.J. and P.A. Yotopoulos, "A Test of Relative Efficiency and Applications to Indian Agriculture", American Economic Review, 61 (1971), 94-109.

Lee, L.F. and W.G. Tyler. "The Stochastic Frontier Production Function and Average Efficiency: An Empirical Analysis". Journal of Econometrics, 7 (1978), 385-390.

Lovell, C.A.K and P. Schmidt "A Comparison of Alternative Approaches to the Measurement of Productive Efficiency", in A. Dogramaci and R. Fare (ed.) Applications of Modern Production Theory, Efficiency and Productivity, Boston: Kluwer Academic Publishers, 1988, Chapter 1.

Maindiratta, A. "Largest Size-Efficiency Scale and Size Efficiencies of Decision-Making Units in Data Envelopment Analysis", Journal of Econometrics, 46 (1990), 57-72.

Marschak, J. and W. Andrews, Jr. "Random Simultaneous Equations and the Theory of Production", Econometrica, 12 (1944), 143-205.

Meeusen, W. and J. van den Broeck. "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error". International Economic Review, 18 (1977), 435-444.

Muysken, J. "Estimation of the Capacity distribution of an Industry: The Swedish Dairy Industry 1964-1973", in F. R. Forsund, M. Hoel and S. Longva (ed.). Production, Multi-sector Growth and Planning. Amsterdam: North-Holland, 1985, 43-62.

Nerlove, M. Estimation and Identification of Cobb-Douglas Production Functions. Chicago: Rand McNally & Company, 1965.

Nikaido, H. Introduction to Sets and Mappings in Modern Economics. Amsterdam: North-Holland Publishing Company, 1970.

Olson, J.A., P. Schmidt and D.M. Waldman. "A Monte Carlo Study of estimators of the Stochastic Frontier Production Function. Journal of Econometrics 13, (1980), 67-82.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling. Numerical Recipes in C: The Art of Scientific Computing. Cambridge: Cambridge University Press, 1988.

Richmond, J. "Estimating the Efficiency of Production", International Economic Review, 15 (1974), 515-521,

Royal Statistical Society. "Discussion on Mr. Farrell's Paper". Journal of the Royal Statistical Society. Series A, (General), Part 3, (1957), 282-290.

Russell, R.R. "Measures of Technical Efficiency", Journal of Economic Theory, 35, (1985), 109-126.

Schmidt, P. "On the Statistical Estimation of Parametric Frontier Production Functions", Review of Economics and Statistics, (1976), 238-239.

Schmidt, P. and R.C. Sickles. "Production Frontier and Panel Data", Journal of Business and Economic Statistics. Vol. 2., No. 4, Oct. (1984), 367-374.

Seiford, L.M. and R.M. Thrall. "Recent Developments in DEA - the Mathematical Programming Approach to Frontier Analysis", Journal of Econometrics, 46 (1990), 7-38.

Sengupta, J.K. "Recent Nonparametric Measures of Productive Efficiency", in J. K. Sengupta and G. K. Kadekodi (ed.) Econometrics of Planning and Efficiency. Dordrecht: Kluwer Academic Publishers, 1988, 169-193.

Sengupta, J. K. "Transformation in Stochastic DEA Models", Journal of Econometrics, 46 (1990), 109-123.

Shephard, R.W. The Theory of Cost and Production Functions.

Princeton, N.J.: Princeton University Press, 1953. Reprinted in 1970.

Silkman, R.H. (ed.). Measuring Efficiency: An Assessment of Data-Envelopment Analysis. Jossey-Bass Incorporated Publishers. 1986.

Stevenson, R.E. "Likelihood Functions for Generalized Stochastic Frontier Estimation", Journal of Econometrics, 13 (1980), 57-66.

Varian, H.R. "The Nonparametric approach to Production Analysis", Econometrica, 52 (1884), 579-699.

Varian, H.R. "Goodness-of-fit in Optimizing Models", Journal of Econometrics, 46 (1990), 125-140.

von Neumann, J. "A Model of General Economic Equilibrium", Review of Economic Studies, 13 (1945).

Zellner, A., J. Kmenta and J. Dreze, "Specification and Estimation of Cobb-Douglas Production Functions", Econometrica, 34 (1966), 784-795.