Finding the Frontier:

Methodological Advances in Data Envelopment Analysis

(HET VINDEN VAN DE GRENSLIJN:
METHODOLOGISCHE VERNIEUWINGEN IN DATA ENVELOPMENT ANALYSIS)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE
ERASMUS UNIVERSITEIT ROTTERDAM OP GEZAG VAN DE
rector magnificus

PROF. DR. P. W. C. AKKERMANS M. A.

EN VOLGENS HET BESLUIT VAN HET COLLEGE VOOR PROMOTIES.
DE OPENBARE VERDEDIGING ZAL PLAATSVINDEN OP

DONDERDAG 2 DECEMEBER 1999 OM 11:00 UUR

DOOR

GERRIT TJEERD POST
GEBOREN TE DEVENTER.

PROMOTIECOMISSIE

PROMOTOR:	PROF. DR. J. SPRONK
OVERIGE LEDEN:	PROF. DR. T. KLOEK
	PROF. DR. IR. R. DEKKER
	PROF. DR. P. MOLYNEUX

UNIVERSITEITSBIBLIOTHEEK
Erasmus Universiteit
Burgemeester Oudlaan 50
3062 PA ROTTERDAM

293
f
57

Contents

INTRODUCTION **1**

1. Introduction 1
2. Efficient Frontier Analysis 2
3. Econometric Frontier Estimation 4
4. Data Envelopment Analysis 5
5. Motivation for this study 7
6. DEA building blocks 8
7. Structure of this study 12
8. Acknowledgements 17

CHAPTER 1: THE STANDARD DEA MODEL **19**

1.1. Introduction 20
1.2. The standard model 21
1.3. Example illustration 25
1.4. Statistical properties 27
1.5. Slack correction 30
1.6. Dual formulation 33
1.7. Concluding remarks 35

CHAPTER 2: ADVANCES IN PREFERENCE MODELLING **37**

2.1. Introduction 38
2.2. Improving goal congruence 40
2.3. Interactive Data Envelopment Analysis 45
2.4. Empirical application 49
2.5. Concluding remarks 56
Appendix 2.1 57

CHAPTER 3: ADVANCES IN PRODUCTION MODELLING **59**
3.1. Introduction 60
3.2. Alternative production assumptions 61
3.3. Violations of convexity 68
3.4. Transconvex DEA 71
3.5. Convex input sets and output sets 76
3.6. Empirical selection procedure 87
3.7. Simulating statistical properties 90
3.8. An empirical investigation into the production technology of bank branches 98
3.9. Concluding remarks 107
Appendix 3.1 109

CHAPTER 4: ADVANCES IN DISTRIBUTION MODELLING **113**
4.1. Introduction 114
4.2. Improving DEA robustness 115
4.3. The general MVDEA framework 120
4.4. Special cases 134
4.5. Emprical application 137
4.6. Monte-Carlo simulations 143
4.7. Concluding remarks 147

5. CONCLUSIONS AND RECOMMENDATIONS **149**
5.1. Introduction 149
5.2. Conclusions 149
5.3. Recommendations 152

REFERENCES **157**

NEDERLANDSTALIGE SAMENVATTING **171**
(SUMMARY IN DUTCH)

Introduction

1. Introduction

Data Envelopment Analysis (DEA) is a methodology for analyzing production processes and for measuring productive efficiency. It originated from the seminal work by Charnes, Cooper and Rhodes (1978), Banker, Charnes and Cooper (1984) and Charnes, Cooper, Seiford and Stutz (1985). In the relatively short span of two decades, DEA has established itself as an important analytical tool. An increasing number of applications are evidence of its popularity among researchers and practitioners in the business community and in government institutions. However, the methodology suffered initially from a number of limitations. Specifically, the original DEA models imposed a number of assumptions that can be restrictive in particular research environments. In addition, it was not clear initially how to test those assumptions and how to alter them. However, a large number of methodological developments have greatly improved its flexibility to deal with a wide variety of problems. This study surveys these methodological developments. In addition, it develops new extensions on the basis of Multi-Criteria Decision-Making theory, the mathematics of Generalised Convexity and the theory of Stochastic Dominance. The purpose of this study is to disseminate knowledge about DEA, its strengths and weaknesses, and highlight the tools currently available for exploiting its full potential.

This introductory chapter has multiple purposes. First, Section 1.2 to 1.4 place DEA in the broader context of the economic literature. DEA is the nonparametric approach to Efficient Frontier Analysis (Section 1.2), henceforth EFA. That is a general methodology for analysing production relationships and productive efficiency based on efficient frontiers rather than central tendencies. That methodology also comprises Econometric Frontier Estimation (Section 1.3), henceforth EFE, which is the parametric approach to EFA. Section 1.4 discusses the distinguishing features of DEA. DEA does not require an explicit specification of the functional form of the production technology and the distribution of inefficiency. Rather, DEA imposes general production and distribution assumptions only. This feature gives DEA a comparative advantage relative to EFE, which typically does require such a specification explicitly. In addition, DEA solves a separate mathematical programming problem for each decision-making unit (DMU) in the data set. The solutions provide detailed information about the nature of the production technology and the efficiency of the DMUs, including efficiency measures, reference units, and shadow prices. A distinguishing feature of DEA is that this information is tailored to each DMU individually. By contrast, EFE optimises a single regression plane that is assumed to apply to each DMU.

Second, Section 1.5 gives a personal motivation for this study. Enthusiasm about the power and the flexibility of the current methodology, as well as dissatisfaction about the fact that this power and flexibility have not been fully appreciated among economists and econometricians, motivate this study. The limitations of the original models are well known. However, the strengths, as well as the recent developments that remedy the limitations of the original models, have received less attention.

Third, Section 1.6 and 1.7 give the structure of this study. The nonparametric approach is not a panacea. The goodness of the results generally depends on the

quality of the data and the goodness of the model used for analysing the data. More specifically, DEA models rely on particular sets of assumptions. All models require certain assumptions, so as to ensure goal congruence and to limit finite sample error (see the discussion below). Therefore, the ability to elicit information and in addition to include information in the mathematical programming model is critical to DEA. The information available on decision-maker preferences, production relationships and data distribution will differ from one research environment to another. At issue is not so much what assumptions should be imposed, but rather how information can be elicited and included in the model, and in addition how to deal with imperfect information. This study discusses the tools currently available for enhancing that ability. Section 1.6 distinguishs 3 types of assumptions of DEA models: assumptions on decision-maker preferences, assumptions on the production technology, and assumptions on the statistical distribution of the observations in the data set. Section 1.7 elaborates on the structure of the study in more detail and gives a brief survey of the contributions developed in the subsequent chapters. Finally, Section 1.8 acknowledges the contribution of people that gave invaluable input to this study.

2. Efficient Frontier Analysis

The analysis of productive efficiency is one of the cornerstones of modern economics[1]. In general, productive efficiency evaluates the achievements of decision-making units[2] (henceforth DMUs) in performing particular production activities. Such activities generally involve the conversion of inputs (e.g. labour and capital) into outputs (e.g. products and services), as displayed in Figure 1. Efficiency concerns the extent to which DMUs succeed in performing the productive activities according to particular organisational goals.

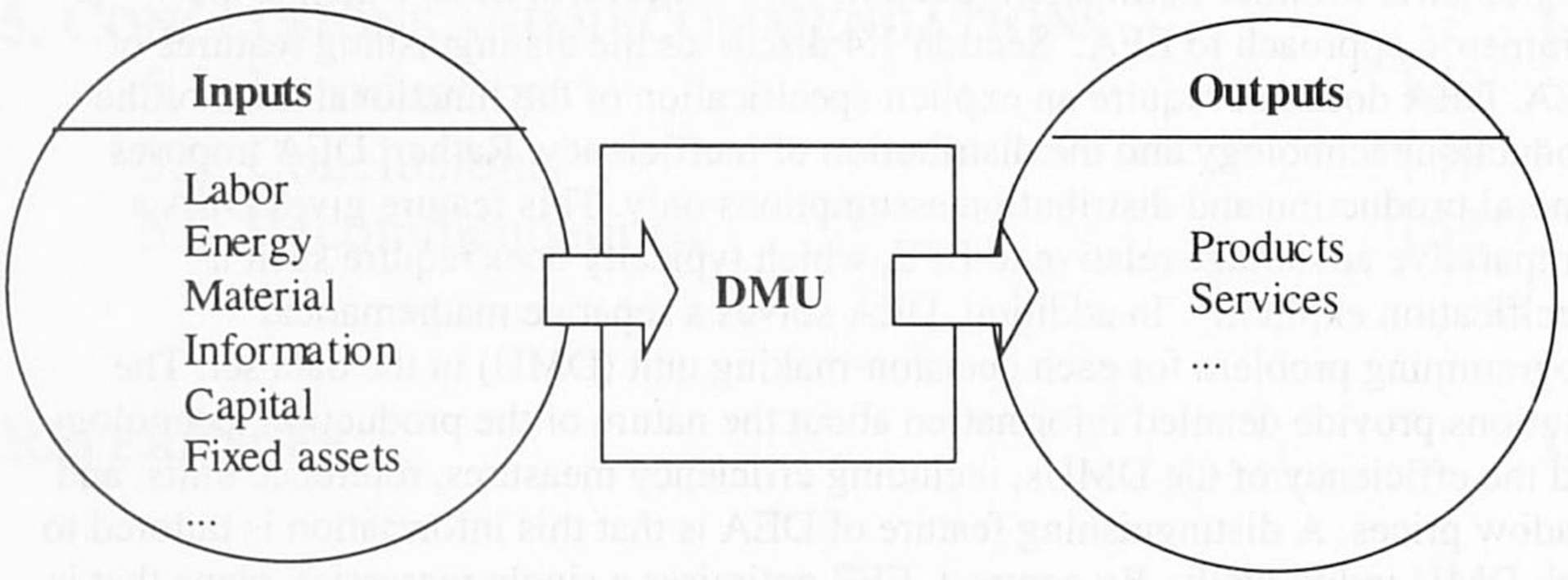


Figure 1 The conversion of inputs into outputs by a DMU

[1] For a comprehensive introduction on this subject, see Färe et al. (1984, 1994) and Coelli et al. (1998).

[2] The term decision-making unit forms part of the DEA idiom. In principle, DMUs can include individual persons, organisations, as well as whole economies. Examples of DMUs in DEA applications include both non-profit organisations such as public schools, university departments and hospitals, and profit organisations such as financial institutions and bank branches. The criterion is that the DMU has control over the input and outputs of the production activities that are evaluated (although nondiscretionary variables can be accommodated as well).

The analysis of efficiency is directly relevant for academic research, public policy-making and managerial decision-making. Farrell (1957) states in his seminal chapter on efficiency measurement:

> *The problem of measuring the productive efficiency of an industry is important to both the economic theorist and the economic policy maker. If the theoretical arguments as to the relative efficiency of different economic systems are to be subjected to empirical testing, it is essential to be able to make some actual measurements of efficiency. Equally, if economic planning is to concern itself with particular industries, it is important to know how far a given industry can increase its outputs by simply increasing its efficiency, without absorbing further resources'.*

A rigorous analytical approach to studying and measuring productive efficiency originated from the seminal work by Debreu (1951), Koopmans (1951, 1957) and Farrell (1957). That approach builds on the neo-classical theory of production. In that theory, production is considered as a problem of optimisation under technological constraints. The technological constraints are represented by an *efficient production frontier*[3]. The efficient frontier represents the maximum amounts of output that can be produced by given amounts of input or, alternatively, the minimum amounts of input required to produce given amounts of output. Inefficiency is defined as the deviation of actual behaviour from 'optimal' behaviour given the production frontier. Unfortunately, production theory can not be applied directly in practice, because the information on the production possibilities generally is incomplete. EFA constructs an empirical frontier on the basis of a data set of comparable DMUs, and measures the efficiency of the DMUs relative to that frontier.

The use of frontier models is becoming increasingly widespread for a variety of reasons. For one thing, frontiers are consistent with the assumption of optimising behaviour that pervades economic theory. Färe, Grosskopf and Lovell (1994, p. 3-4) cite 3 additional motives for estimating efficient frontiers and measuring efficiency relative to the estimated frontiers:

> *'(a) The structure of economic frontiers is of inherent interest. It is important to know whether, for example, technologically efficient production is characterised by economies of scale or scope, and whether cost efficient production is characterised by subadditivity.*
>
> *(b) The structure of economic frontiers can be different from the structure of economic functions constructed from the same data. Best practice is not just better than average practice, it may also be structurally different. It is important to know whether, and if so in what ways, the structure of efficient production differs from the structure of average production. Best practice may be better than average practice precisely because it exploits available substitution*

[3] It is important to distinguish frontier or extremal relationships from 'central tendency' or 'average' relationships. Empirical production analysis has long focused on central tendency. However, theoretical production analysis has always focused on frontiers, recognising that production activities represent a constrained optimisation problem. Therefore, efficient frontier analysis bridges the gap between theoretical and empirical research of producer behaviour.

possibilities or scale opportunities that average practice does not. Public policy based on the structure of best practice frontiers may be very different from policy based on the structure of average practice.

(c) The distance between production frontiers and observed producers is of obvious policy interest. It is important to know how inefficient observed production is on average, and at what cost. It is also important to know what types of producers are most efficient, and what type of inefficiency is most prevalent.'

Unfortunately, information on the true frontier typically is incomplete. However, for practical application, an empirical approximation of the frontier can be obtained from a data set of empirical production data. In the last two decades various elaborate ways of estimating frontiers and measuring efficiency have been developed. Two competing methodologies can be distinguished in the literature, frequently labelled *Econometric Frontier Estimation*, or the parametric approach, and *Data Envelopment Analysis*, or the non-parametric approach.

3. Econometric Frontier Estimation

One empirical approach to efficient frontier analysis is Econometric Frontier Estimation, henceforth EFE. The econometric approach uses regression analysis to estimate frontiers. Specifically, EFE first specifies a regression equation along with a two-part composed error term. One part of the composed error term represents inefficiency and is typically assumed to follow a particular one-sided distribution, such as the truncated normal distribution used in Stevenson (1980)[4]. The other part represents statistical noise and is generally assumed to follow a normal distribution. The regression equation fixes the functional form (e.g. linear or log-linear) of the production frontier[5], but the parameters are left unspecified. EFE next selects the parameters that make the regression equation fit the data best by optimising a particular statistical goodness measure (e.g. the sum of squared errors or a likelihood function). Aigner, Lovell and Schmidt (1977), Meeusen and Van den Broeck (1977) and Battese and Corra (1977) first proposed this approach. Finally, observation specific estimates of inefficiency can be obtained by using the distribution of inefficiency conditional on the estimate of the entire composed error term, This approach was proposed first by Jondrow, Lovell, Materov and Schmidt (1982) and Kalirajan and Flinn (1983). Schmidt (1985-86), Lovell and Schmidt (1988) and Bauer (1990) provide an excellent introduction to the EFE literature.

An important limitation of EFE is that it typically requires the specification of an explicit, and possibly overly restrictive, functional form for the production relationships, as well as an explicit distribution for the inefficiencies. This requirement can reduce the goodness of the econometric approach if meaningful theoretical foundations and reliable statistical tests for the functional form and the inefficiency distribution are absent.

---

[4] However, if panel data are available, the specification of a particular distribution can be avoided by assuming a particular time series pattern for inefficiency. For more details, see Schmidt and Sickles (1984), Cornwell, Schmidt and Sickles (1990), and Hay and Liu (1997).

[5] However, within the parametric approach, flexible functional forms have been developed that offer substantial flexibility (e.g. Diewert and Wales (1987) and Pulley and Braunstein (1992)). For example, the well-known translog estimation function can give a second-order approximation to an arbitrary twice-differentiable function.

In addition, EFE focuses on obtaining a global representation of the production technology rather than the local production possibilities of specific DMUs. Sometimes it is desirable to obtain detailed DMU-specific information on production possibilities and efficiency, especially in managerial applications. Moreover, it is generally desirable to find out causes and remedies for inefficiencies in addition to measuring them.

4. Data Envelopment Analysis

Data Envelopment Analysis, henceforth DEA, is an alternative methodology for efficient frontier analysis. It has been developed in the field of Operations Research/ Management Science and uses mathematical programming techniques rather than regression techniques.

DEA evolved over the past 20 years, starting with the Ph. D. dissertation of Edwardo Rhodes (1978) under the supervision of professor W. W. Cooper. His research involved estimating the relative efficiency of different public schools that either did or did not participate in a special educational program for disadvantaged students, called Program Follow-Through. That research resulted in the formulation of the Charnes, Cooper and Rhodes (1978) model. Hence, DEA began as an OR/MS tool for technical-efficiency analysis of public-sector DMUs. Since the formulation of the CCR model, there has been much methodological development and empirical application of DEA. For example, Seiford (1996) lists over 700 DEA articles, studies and dissertations, and many more have appeared since. The bibliography is evidence of DEA applications in a wide range of contexts including education, health care and banking. For an excellent introduction in the DEA methodology, see Charnes, Cooper, Lewin and Seiford (1994).

The distinguishing feature of DEA is the absence of strong assumption about the functional form of the production relationships and the distribution of the inefficiencies, i.e. its nonparametric nature. Rather than imposing an explicit parametric specification, DEA imposes general production and distribution assumptions only. This feature makes it possible for the data to 'speak for themselves' rather than being forced to use the idiom of some imposed functional form. This gives DEA a comparative advantage relative to econometric models. Bowlin, Charnes, Cooper and Sherman (1985), Banker, Charnes, Cooper and Maindiratta (1988), Gong and Sickles (1989, 1992), Banker, Gath and Gorr (1993), and Bojanic, Caudill and Ford (1998), among others, provide comparisons of DEA and EFE based on artificial data. In addition, Banker, Conrad and Strauss (1986) and Ferrier and Lovell (1990), among others, compared DEA and EFE using real-life data. The majority of empirical production analysis in Operations Research/Management Science uses DEA, while most empirical analysis in economics uses EFE[6].

However, DEA need not be seen as a competitor of EFE. Rather, DEA can provide new insights and information in addition to that available from econometric

---

[6] However, there is a substantial overlap with the literature on nonparametric production analysis in economics, as developed by Afriat (1972), Hanoch and Rothschild (1972), Diewert and Parkan (1983), Varian (1984, 1985), Banker and Maindiratta (1988), among others. That literature uses mathematical programming models that are similar to DEA models. Initially, these models dealt with testing regularity conditions and behavioural objectives, for the purpose of screening data for accuracy and outliers, and guidance in the selection of parametric functional forms to use in econometric estimation. However, recently it has been proposed to use these models to construct frontiers and measure efficiency of data relative to the constructed frontiers, bringing this line of research closer to DEA. A similar line of research has been developed for the nonparametric analysis of consumer behaviour, including Afriat (1967), Diewert (1973), Diewert and Parkan (1978), and Varian (1982).

methods. In this respect, another distinguishing feature of DEA is its focus on obtaining DMU-specific information. In contrast to parametric approaches, the objective of which is to optimise a single regression plane that is assumed to apply to each DMU, DEA gives information that is tailed to each DMU individually. For each DMU in the data set, DEA solves a separate mathematical programming problem. DEA solutions provide detailed information about the nature of the local production possibilities and the efficiency of the DMUs. This study distinguishs three types of output generated by DEA models:

(1) *efficiency measures*,

(2) *reference units*, and

(3) shadow prices.

First, DEA models compute efficiency measures. In general, these measures are constructed as measures of distance from the empirical frontier. The efficiency measures can be used as a summary measure for the efficiency of past operations and for assessing potential improvements for future operations. Hence, they provide valuable information for managerial planning and control[7]. In addition, academic researchers can use the efficiency scores to test economic theories.

Second, DEA identifies reference units, which are observed DMUs, and possibly fictitious DMUs, constructed as combinations of observed DMUs. The reference units represent revealed best practice. The referencing input-output levels are potential performance targets. In addition, the reference units constitute potential benchmark partners. Comparing the work process of inefficient DMUs with that of these units can reveal causes and remedies for inefficiencies.

Finally, DEA models compute virtual weights for the inputs and outputs. The weights are computed as the marginal production possibilities (for the inputs) and the marginal input requirements (for outputs) at the reference point on the empirical frontier. The virtual weights can be used for measuring local scale, substitution and transformation characteristics of the production technology, and hence provide valuable information on the production technology for production planning and economic analysis.

---

[7] For a discussion of DEA as a tool for managerial control and diagnosis, see Epstein and Henderson (1989).

5. Motivation for this study

Both enthusiasm and frustration motivate this study. When I was first introduced to DEA some four years ago, I adopted a rather skeptical attitude. Working at Tinbergen Institute, a research institute at which 'traditional' Econometrics prevails, I suppose that at least some skepticism was to be expected. After all, DEA originated from the world of Operations Research/Management Science, a world (at least in my perception) inhabited by 'deterministic' people rather than the 'stochastic' people that inhabit the world of Econometrics. In addition, no-one at the institute had ever heard about DEA, except for one of the directors, who always looked somewhat discouraged when I spoke of the subject, and who knew of chapters in econometric top-journals that hammered DEA already many years ago. All other people simply associated the three-letter acronym with the Drug Enforcement Agency.

Nevertheless, I decided to give DEA the benefit of the doubt, and dug into the literature, admittedly after my supervisor kept pushing the subject. After reading the Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984) classics, I started to realize that DEA actually had much to say for it. I was especially excited about its promise of freeing the investigator from having to impose prior structure on economic relationships of interest. Finally, so it seemed, an analytical tool was available that allowed data to speak directly for themselves, without being forced in the idiom of the preconceived opinions of the researcher. An objective tool that would overcome the shortsightedness of the established techniques, and enable researchers to objectively analyze economic reality. At some point, I even saw a parallel with the Hubble Space Telescope, which is said to give 'mankind's deepest, most detailed optical view of the universe, unveiling bewildering galaxies across billions of years'[8].

However, I also recognized some of the limitations inherent in the earliest DEA models. More specifically, notwithstanding its nonparametric nature, DEA models impose particular assumptions. Those assumptions involve the preferences of the decision-maker, the nature of the production technology, and the statistical distribution of the observations in the data set. All models require particular assumptions, so as to ensure goal congruence of the results and to avoid finite sample error (see Section 6 for more details). In addition, the information available differs from one research environment to another. Therefore, it is desirable to have the flexibility to alter the assumptions to closely match the research environment. The original models impose particular assumptions. Those assumptions fit certain problems, but may be too restrictive for other problems, including those with high preference ambiguity, non-convex production characteristics and stochastic disturbances. In addition, it was not clear initially how those assumptions could be tested or altered. This limitation made me think about Hubble once again, as a flaw in the telescope's main mirror had significantly reduced its ability to focus, giving blurred images due to spherical aberration, since its launch in 1990.

However, as I became more familiar with the DEA literature, I found out that a number of extensions had been developed since the seminal CCR and BCC articles that can help remedy the limitations of the original models. These extensions include the development of a wide variety of alternative models, each imposing a different set of assumptions and fit for a different research environment, the use of Multi-Criteria Decision making techniques that can help to elicit preferences from the decision-

---

[8] See http://www.stsci.edu/.

maker, empirical specification tests to guide in the selection of the model that fits a particular data set best, and outlier detection procedures and stochastic DEA models.

These extensions made me think of the corrective optics package carried out by the 1993 Endeavour mission restored Hubble's intended imaging capabilities. I strongly believe that DEA is one of the most powerful and flexible analytical tools currently available to analysts and practitioners, capable of dealing with a wide variety of different problems.

The shortcomings of DEA are fairly well known among economists and econometricians. However, the benefits of DEA as well as the current advances that circumvent some of the important limitations inherent in the earliest models have not been fully appreciated. The enthusiasm about the current state of the DEA methodology and the frustration about the ignorance about this state motivated me to write this study. It surveys some of the important methodological advances in the DEA literature, and presents some new contributions. I hope this study will contribute to the dissemination DEA, the knowledge of its relative strengths and weaknesses, and the tools currently available for exploiting its full potential.

6. DEA building blocks

Its nonparametric nature and its focus on individual DMUs give DEA a comparative advantage relative to econometric techniques. Still, DEA is not a panacea. As is true for most empirical techniques, the goodness of the DEA output depends on the quality of the data set. Reliable results generally require a rich data set that contains many comparable DMUs and that includes many of the relevant inputs and outputs, and that is only minimally perturbed by measurement error. The quality of the data generally is more important in DEA than in the econometric approach, because, as discussed above, DEA lets the data speak for themselves and does not rely on a prior functional form for the frontier and the inefficiency distribution.

An equally important issue is the goodness of the model used for analysing the data. Different models use different assumptions. This study distinguishs three types of assumptions in DEA models:

(1) *preference assumptions*,

(2) *production assumptions*, and

(3) *distribution assumptions*.

First, different models impose different assumptions about decision-maker preferences. Models that focus on technical efficiency exclusively are 'value-free' to some extent. However, performance evaluation generally deals with evaluating effectiveness, i.e. the extent to which DMUs succeed in achieving their organisational goals, rather than technical efficiency only. For example, in many organisations, cost efficiency and profitability are important objectives in addition to technical efficiency. It is important to use models that correctly reflect the organisational objectives of the DMUs (goal congruence), so as to obtain results (efficiency measures, reference units and shadow prices) that are relevant for policy making. For example, a reference unit that is technically efficient need not represent a desirable benchmark if profit maximisation is the main objective.

In this study, we represent preferences by a utility function that is defined over the inputs and outputs[9]. For example, Figure 2 displays three indifference curves for a hypothetical utility function that is defined over a single input and a single output.

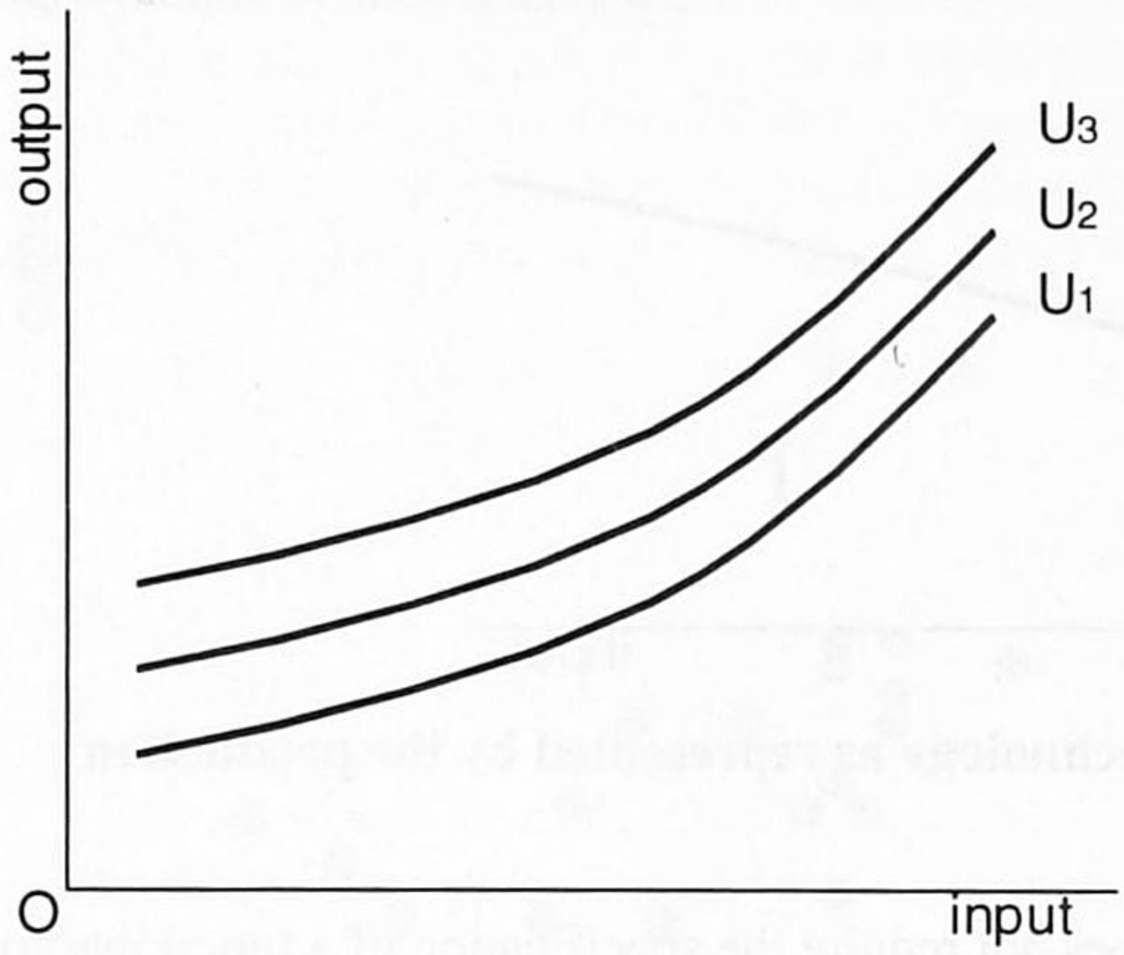


Figure 2 Decision-maker preferences as represented by indifference curves of a utility function.

The reliable elicitation of information on decision-maker preferences and the incorporation of that information in the mathematical programming problem are one of the critical success factors in DEA. Fortunately, in the last two decades, a number of advances in DEA have greatly extended the modelling options. Advances include the use of behavioural assumptions from micro-economic theory and Multi Criteria Decision-Making techniques that can aid in the elicitation of preferences from the decision-maker.
These extensions give DEA a comparative advantage relative to econometric frontier estimation, which does not have much flexibility to include organisational objectives other than efficiency.

Second, different models impose different assumptions about the production technology. Those assumptions are required to build an empirical production set from the observations in the data set. In general, the production technology can be represented by a production possibilities set, which represents all attainable combinations of input and output. For example, Figure 3 displays the production set *T* for a hypothetical single-input single-output technology. The solid line represents the efficient frontier.

---

[9] Note that utility functions usually represent consumer preferences over consumption goods (or attributes of consumption goods) in neo-classical microeconomics. By contrast, producers are usually assumed to maximise the (expected) profit associated with inputs and outputs, rather than to maximise the (expected) value of a utility function. However, in principle nothing excludes the application of utility analysis to producer behaviour, provided the utility assumptions for consumers are not adopted gratuitously for producers. Moreover, the traditional profit function can be interpreted as a special case of utility function. A number of authors have used utility functions to represent DMU preferences over inputs and outputs in DEA, including Halme *et al.* (1999).

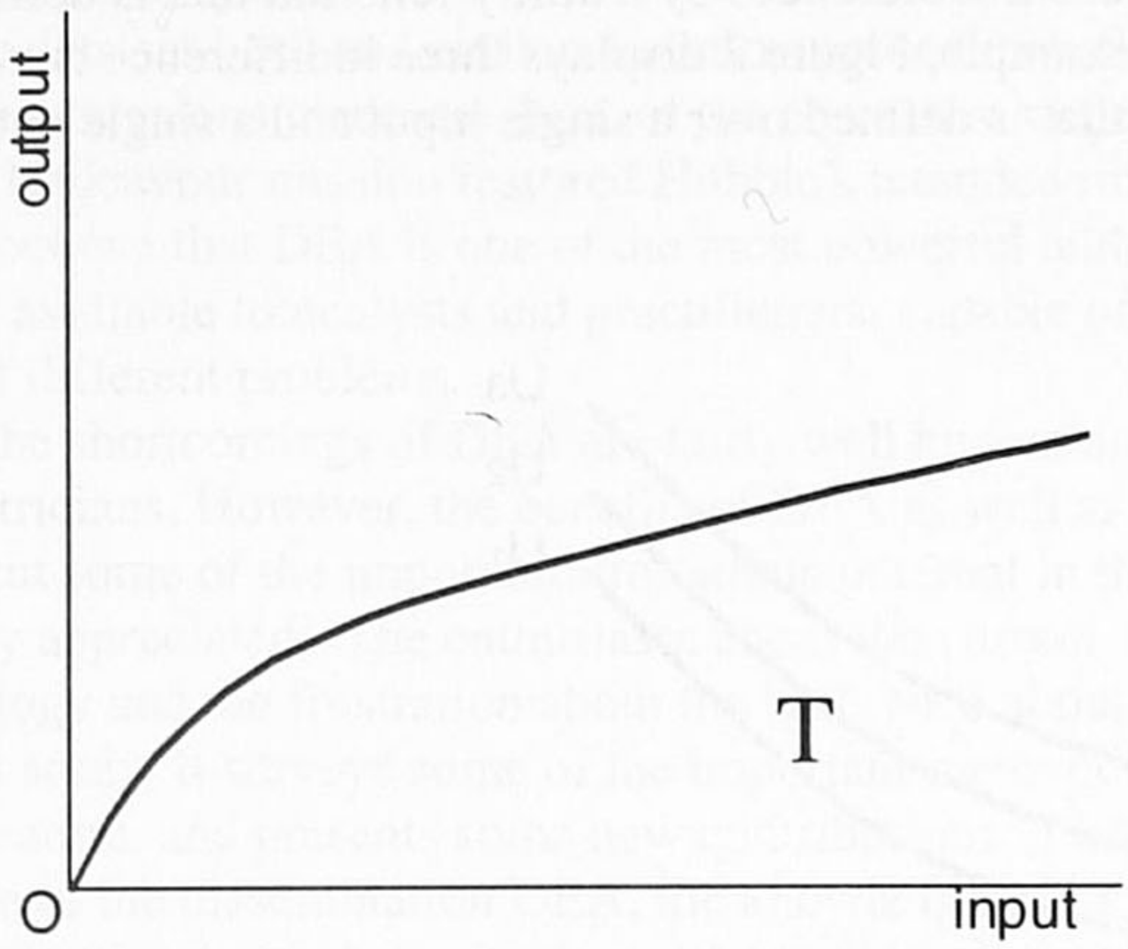


Figure 3 The production technology as represented by the production possibilities set *T*.

As discussed above, DEA does not require the specification of a functional form for the production relationships. Rather, DEA relies on assumptions about the general nature of the production technology. This gives DEA a comparative advantage relative to econometric frontier estimation, which typically do require an explicit functional form.

In fact, certain DEA models impose minimal production assumptions only. For example, the Free Disposable Hull model (Deprins, Simar and Tulkens (1984)) assumes free disposability[10] for the production possibilities only. Unfortunately, models that impose minimal prior structure have difficulty in discriminating between efficient and inefficient DMUs, especially if the data set contains relatively few efficient observations relative to the number of input-output variables (finite sample error). This can substantially reduce the goodness of DEA output. Efficiency measures may be biased, reference units may give minimal clues for improving efficiency, and the input-output weights may not reflect the marginal properties of the technology.

Incorporating additional information about the production technology can improve small sample performance. For example, the standard Banker, Charnes and Cooper (1984) model assumes convexity for the production possibilities in addition to free disposability. Introducing such information can substantially improve the analysis. The more structure we impose on a model the better the estimates, provided the structure is correct. Unfortunately, imposing additional assumptions potentially introduces mis-specifications. For example, assuming convexity can make small organisations look spuriously inefficient relative to large reference units, in case of increasing returns-to-scale. Therefore, the reliable elicitation of production information and the incorporation of that information in the model are critical to the success of DEA assessments.

---

[10] A technology satisfies free disposability if increasing input does not reduce the production possibilities, and reducing output does not increase the input requirement. It corresponds to the typical micro-economic assumption that the marginal productivity of inputs is non-negative. For more details on the production assumptions imposed by different DEA models, see Chapter 1 and Chapter 3.

Finally, different models impose different assumptions on the statistical distribution of the observations in the data set. To determine what the DMUs have actually achieved, and in addition to construct the empirical production set that the achievements are compared with, DEA uses a data set of comparable DMUs. Figure 4 displays an example of such a data set.

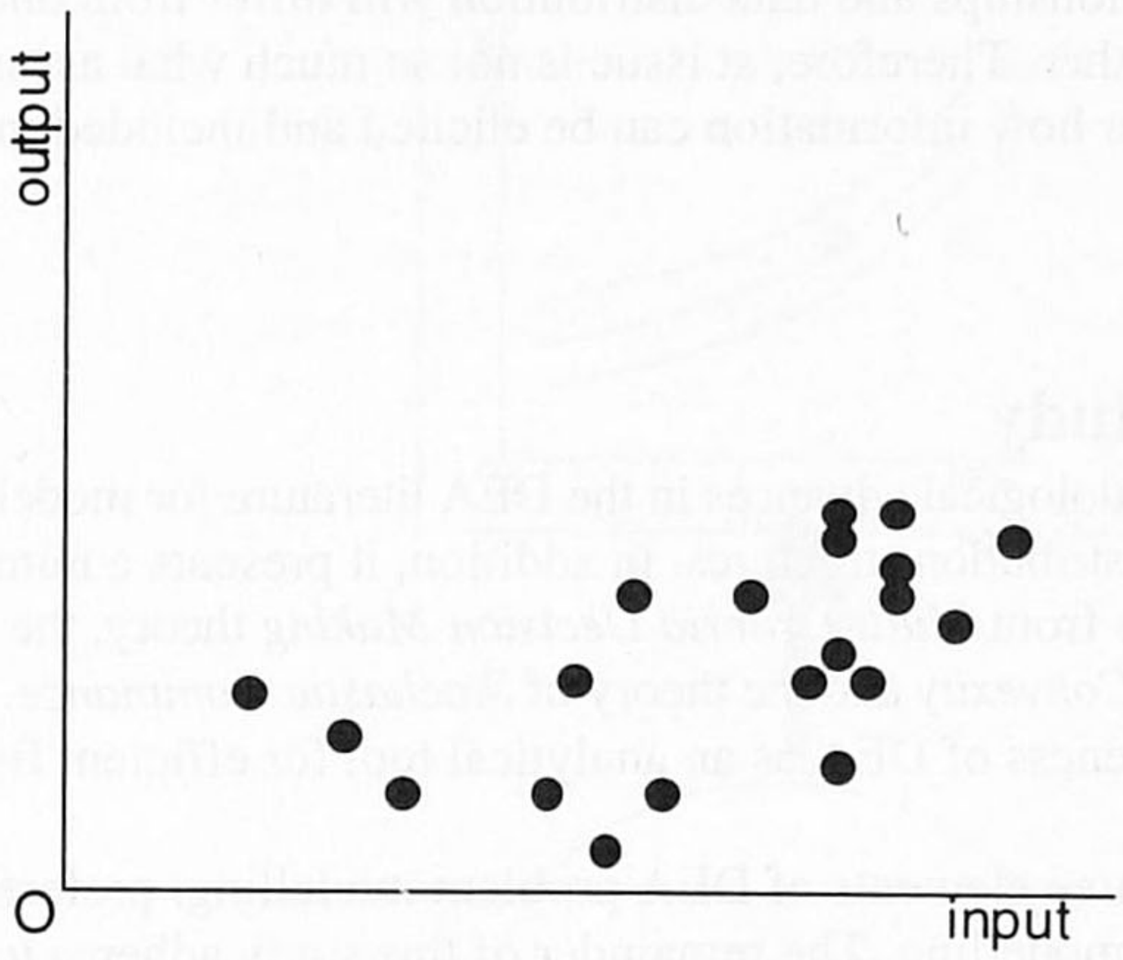


Figure 4 Example of a data set used for approximating the production possibilities

As discussed above, DEA models generally do not impose a particular functional form for the distribution of the inefficiencies. Rather DEA relies on general assumptions about the distribution of the observations. The original DEA models assumes that the DMUs in the data set represent a sample from the production set. In addition, it is assumed that the input-output vectors of the DMUs are measured accurately. Therefore, these models are deterministic models. By contrast, econometric frontier estimation does systematically account for stochastic disturbances.

These assumptions can be violated by *DMU heterogeneity* and *parameter uncertainty*. DMU heterogeneity occurs if the DMUs operate under different technologies, and hence face different production sets. Another problem is parameter uncertainty for the input-output vectors of the evaluated DMUs. Traditional DEA neglects parameter uncertainty by treating the observations as if they were the true values for the input-output variables. Stochastic disturbances such as measurement error, random noise, outlier observations or external effects may violate this assumption. In stochastic environments, the input-output variables are unknown and have to be estimated using prior theory or available economic data. In many cases, theory and data allow for an inaccurate estimation only, and there is uncertainty regarding the true input-output variables.

Such violations can seriously reduce the value of DEA output. The emphasis of EFE rather than DEA among economists is due in large part to the fact that the standard models do not account for error in addition to inefficiency, and therefore can be contamined with outliers and measurement error (although EFE is also sensitive to outliers and in addition to specification error). However, recently a number of extensions have been developed that do account for DMU heterogeneity and parameter uncertainty. The extensions include the use of outlier detection procedures

and Stochastic Programming techniques. Unfortunately, these techniques can reduce the discriminating power of DEA assessments. Therefore, the reliable elicitation of distribution information and the incorporation of that information in the model are one of the critical success factors in DEA.

No model can do without assumptions as that endangers goal congruence and introduces finite sample error. In addition, the information available on decision-maker preferences, production relationships and data distribution will differ from one research environment to another. Therefore, at issue is not so much what assumptions should be imposed, but rather how information can be elicited and included in the model.

7. Structure of this study

This study surveys the methodological advances in the DEA literature for modelling preference, production and distribution structures. In addition, it presents a number of new advances, using insights from *Multi-Criteria Decision-Making* theory, the mathematics of *Generalised Convexity* and the theory of *Stochastic Dominance*. These advances extend the effectiveness of DEA as an analytical tool for efficient frontier analysis.

Figure 4 shows the three elements of DEA problem modelling: preference, distribution, and production modelling. The remainder of this study adheres to this structure.

Chapter 1: The standard DEA model

Chapter 1 provides the foundation for the chapters that follow. It discusses the basic concepts and characteristics of the DEA methodology, and in addition introduces the notation that will be used throughout the study. For simplicity, Chapter 1 focuses on the output-oriented version of the standard Banker, Charnes and Cooper (1984) model. That model imposes particular production, distribution and preference assumptions. Those assumptions fit certain problems, but may be too restrictive for other problems, including those with high preference ambiguity, non-convex production characteristics and stochastic disturbances. In the subsequent chapters, the possibilities for introducing alternative or additional assumptions will be discussed in terms of this model. For non-experts, reading Chapter 1 is a prerequisite for comprehension of the developments discussed in the subsequent chapters. However, the latter chapters can be read independently. This feature allows readers to take the most efficient route to specific subjects of interest.

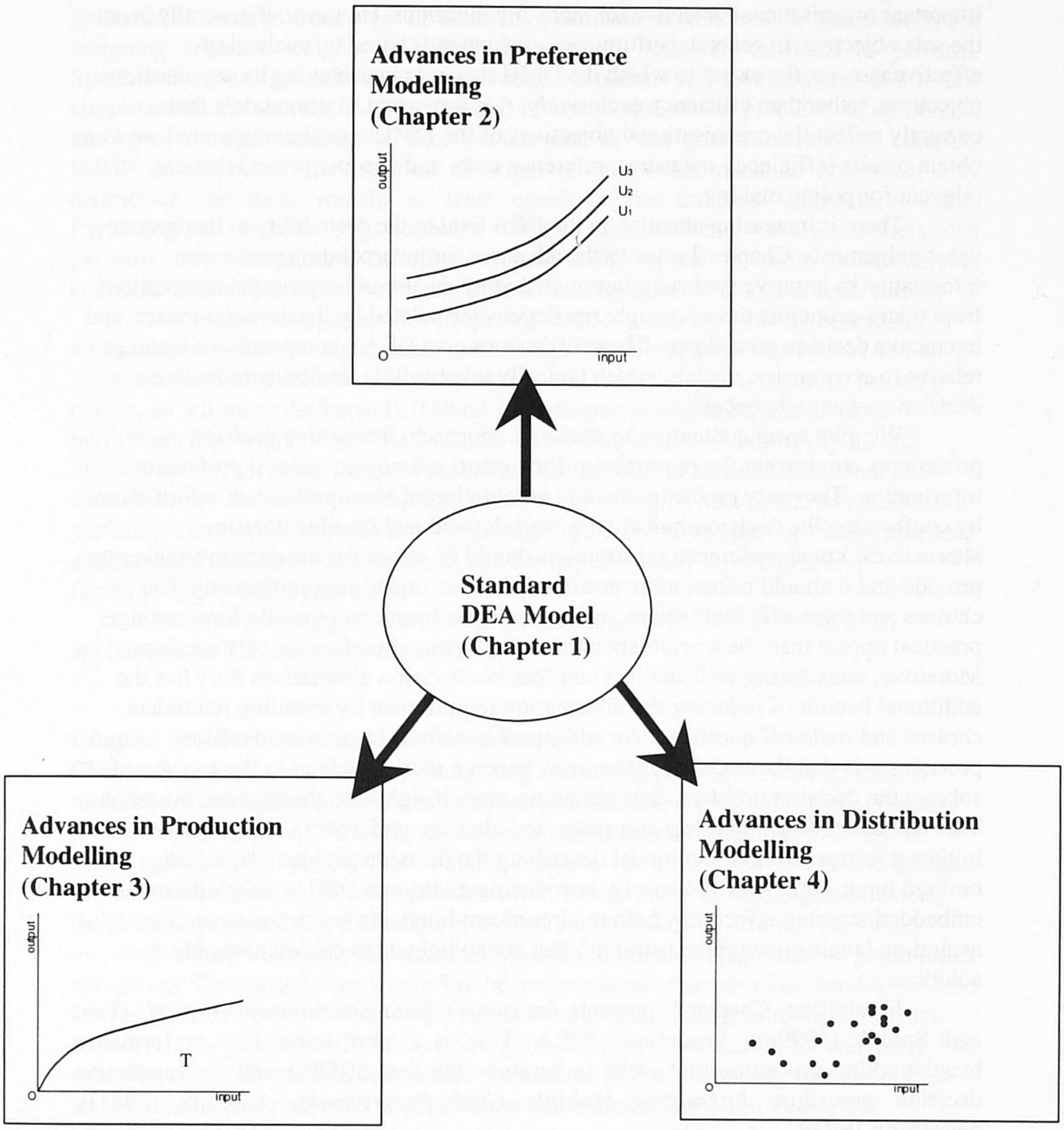


Figure 4 Main structure of the study. After an introductory chapter, three chapters discuss the methodological advances for modelling preference, production and distribution structures.

Chapter 2: Advances in preference modelling

The original DEA models focus on technical efficiency. Technical efficiency is an important organisational objective for many organisations. However, it generally is not the sole objective. In general, performance evaluation is based on evaluating effectiveness, i.e. the extent to which the DMU succeeds in achieving its organisational objectives, rather than efficiency exclusively. It is important to use models that correctly reflect the organisational objectives of the DMUs (goal congruence), so as to obtain results (efficiency measures, reference units and shadow prices) that are relevant for policy making.

There is increasing attention in the DEA field to the desirability to incorporate value judgements. Chapter 2 reviews the advances for incorporating preference information to improve goal congruence, including the use of behavioural assumptions from micro-economic theory, weight restrictions formulated by the decision-maker, and interactive decision procedures. These extensions give DEA a comparative advantage relative to econometric models, which typically involve little flexibility to include decision-maker preferences[11].

We give special attention to the latter approach. Interactive decision procedures circumvent the requirement for a priori articulated general preference information. They rely on the interactive articulation of local preference information, by confronting the decision-maker with well-defined and feasible decision alternatives. Local preference information should be easier for the decision-maker to provide and it should reflect more accurately his/her underlying judgement. The choices and trade-offs involved in answering these questions typically have stronger practical appeal than the a priori specification of general preference relationships. Moreover, considering well-defined and feasible decision alternatives only has the additional benefit of reducing the information requirement by avoiding redundant choices and trade-off questions. An additional benefit of interactive decision procedures is that the decision-maker may become more involved in the process of solving the decision problem, thus obtaining more insight into the decision situation. This embedded learning effect can shape and alter his preferences and may even initiate a restructuring of the model describing the decision problem, by adding omitted input-output variables or by introducing additional DMUs. In addition, the embedded learning effect may help to circumvent imposing restrictions (such as aspiration levels or weight constraints) that are so tight as to cause unfeasible solutions.

In addition, Chapter 2 presents *Interactive Data Envelopment Analysis* (Post and Spronk (1999a)), henceforth IDEA. That is a new interactive performance benchmarking procedure that seeks to combine the best of DEA and the interactive decision procedure *Interactive Multiple Goal Programming* (Spronk (1981)), henceforth IMGP.

Although the proposed interactive techniques circumvent the requirement for a priori articulated general preference information, it is unclear whether the interactive decision procedures they embed can reliably elicit preference information from decision makers in complex decision problems (i.e. problems with large numbers of decision alternatives and goal variables). The fact is, empirical results from experimental

---

[11] Micro-economic assumptions can be included in Econometric Frontier Estimation by estimating cost, revenue, or profit frontiers rather than production frontiers. However, apart from that, EFE involves minimal possibilities for including preference information. For example, EFE typically involves just a single efficiency term that is assumed to apply to the dependent variable (output, cost, revenue or profit).

psychological research (e.g. Larichev *et al.* (1987)) indicate that interactive decision procedures that rely on the information processing operations of selecting a most preferred alternative from a set of alternative decision alternatives, or alternatively on assigning criterion weights or weight bounds, are relatively difficult to perform. Specifically, these procedures can cause decision-makers to be inconsistent and to use simplified heuristics (e.g. to drop criteria) in complex decision problem, i.e. problems that involve large numbers of decision alternatives and goal variables. By contrast, IMGP relies on the identification of criterion values that must be improved, may deteriorate, or must remain at least equal to the attained satisfactory value. Psychological research (e.g. Larichev (1987)) indicates that decision-makers can perform that elementary information processing operation with inconsiderable contradictions and complex strategies (e.g. estimating a set of several criteria values).

An additional argument for proposing IMGP is that the existing interactive approaches consider efficient frontier units only. Therefore, these approaches do not allow the decision-maker to select a reference unit from the PPS interior, if, for any reason, he wishes to deliberately include some degree of organisational slack in performance benchmarking. The decision-maker may wish to intentionally include internal slacks, because of organisational and psychological considerations that are excluded from the original model formulation. For example, best observed practice is a performance standard that can not be exceeded (given observed past operating practice and imposed production assumptions, and consequently may be stressful and demoralising to DMU managers.

To illustrate the operation of IDEA, the Chapter presents the results of an application for UK university departments.

Chapter 3: Advances in production modelling

Chapter 3 reviews the advances for imposing and testing production assumptions. A multitude of alternative model formulations exists that allow for a variety of alternative production assumptions. In addition, empirical specification tests can help select the specification that fits the data best. Finally, nonobserved DMUs can be added to the data set to include additional, local production information. These advances greatly extend the flexibility of the DEA methodology to deal with a variety of different problems.

Still, most DEA models maintain the assumption that the production possibilities are convex. That assumption simplifies the mathematical structure of the model, because convex problems can be solved with standard mathematical programming techniques. However, it is generally difficult to find a good economic justification for convexity. For example, convexity excludes economic phenomena such as increasing returns-to-scale and increasing marginal rates of substitution and transformation can violate this assumption. Therefore, it is desirable to develop non-convex models.

Chapter 3 presents such a model, *Transconvex Data Envelopment Analysis* (Post (1997b), (1998), Dekker and Post (1999)), henceforth TDEA. That model uses some key insights from the mathematical theory of *Generalised Convexity*. Specifically, TDEA replaces convexity with the more general assumption that the production set is convex transformable, or transconvex, using particular transformations for the input-output variables. Although the TDEA approach can accommodate non-convex production possibilities, the mathematical problem remains convex, and can be solved using standard mathematical programming techniques.

Sometimes economic theory can be used to select the appropriate convexifying transformations. For example, microeconomic theory typically assumes that the input

and output isoquants of the production set are convex. Petersen (1990) proposed to use this assumption in DEA rather than the more restrictive assumption that the entire production set is convex. Unfortunately, the model Petersen proposed to do so suffers demonstrably from a number of serious limitations. However, particular exponential transformations can impose the Petersen assumption in the TDEA approach.

Apart from economic theory, empirical selection procedures can provide guidance for selecting the appropriate transformations. We develop a specification test based on exponential transformations and the *F*-test by Banker (1993). Interestingly, that procedure induces a gradual transition between the FDH and BCC models. Those models represent the limiting cases of the exponential transformation.

Apart from theoretical development, the Chapter presents the results of Monte-Carlo simulations, and an empirical application for the bank branch network of a large Dutch financial institution.

Chapter 4: Advances in distribution modelling

The original DEA models assume that the DMUs in the data set represent a sample of a common production set, and in addition that the input-output variables can be measured accurately. chapter 4 reviews the advances presented in the literature for dealing with DMU heterogeneity and parameter uncertainty, including outlier detection procedures and Stochastic Programming techniques. In addition, it presents *Mean-Variance Data Envelopment Analysis* (Post (1997a)), henceforth MVDEA, a novel framework for incorporating parameter uncertainty.

MVDEA considers the selection of reference units in case of parameter uncertainty as a special case of choice problem under uncertainty, and applies the theory of *Stochastic Dominance* to that problem. That theory does not require a specific functional form for the preferences of the decision-maker or a specific distribution function for the statistical distribution of the choice variables. Therefore, Stochastic Dominance it is consistent with the nonparametric nature of the original DEA framework. By contrast, the existing stochastic DEA models invariably require the specification of a particular statistical distribution function for the disturbances.

The original models are consistent with stochastic dominance only if the input-output variables are specified accurately or alternatively if the evaluator is risk neutral with respect to parameter uncertainty. If the input-output variables are uncertain and the evaluator is risk averse, the original models may include reference units that are stochastically dominated and exclude units that are stochastically nondominated. Particular variance corrections can include all stochastically nondominated units, and exclude all stochastically dominated units. An empirical application for large Indonesian banks will demonstrate that the associated reduction in parameter uncertainty can be substantial. In addition, particular variance restrictions can improve DEA efficiency estimates. Monte-Carlo simulations suggest that the improvement can be substantial, especially if a large sample is available and the sample is heavily perturbed by error.

Chapter 5: Conclusions and recommendations

Chapter 5 summarizes the main conclusions of the previous chapters and offers some recommendations for future research. Section 5.2 summarizes the main conclusions. As discussed above, there is much reason for enthusiasm about the methodological advances in DEA. These advances greatly improve the power and the flexibility of DEA to generate valuable information in many different research environments. However,

some problems remain unresolved. Section 5.3 discusses the unresolved problems that were most relevant in performing this study. The discussion is intended as a marginal note for practical applications of the current DEA methodology. In addition, the remarks are intended as a call for future methodological research.

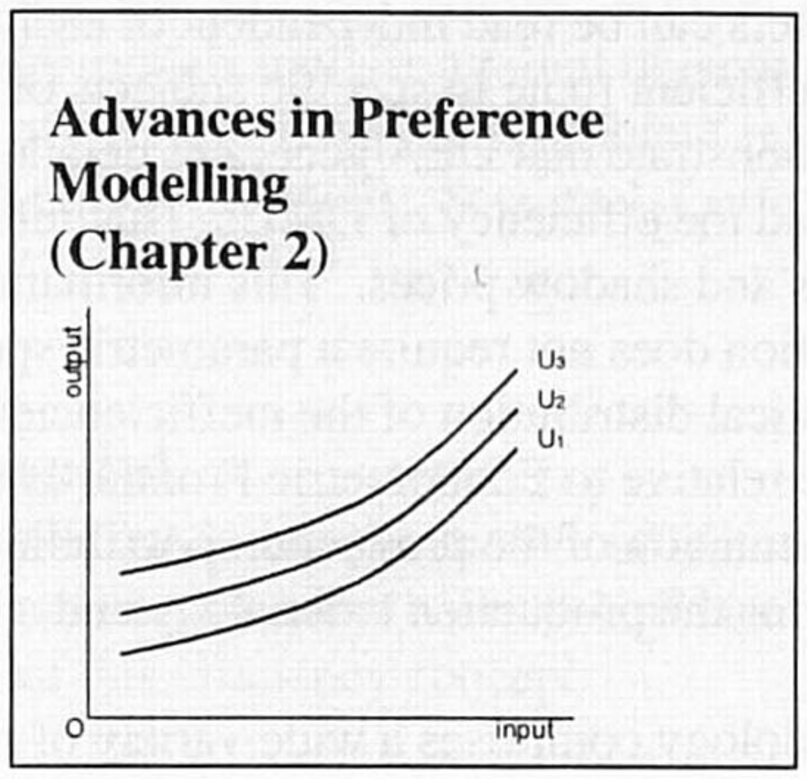
8. Acknowledgements

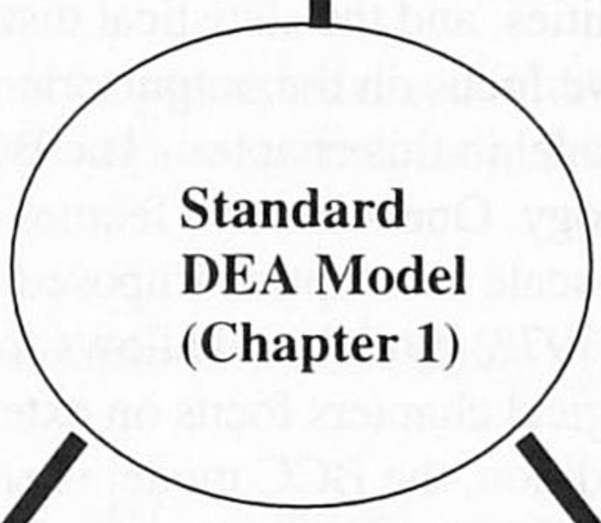
I would like to express my appreciation to several people that gave invaluable input to this study. First, Prof. Jaap Spronk deserves much credit for providing me with many intellectual challenges, the freedom to freely select from them, and the facilities to work on them at ease. I truly believe research is done best at the time (during drowsy office hours or with an unexpected adrenaline rush in the middle of the night) and the place (in the office, on the balcony at home, or for that matter under the shower) of convenience.

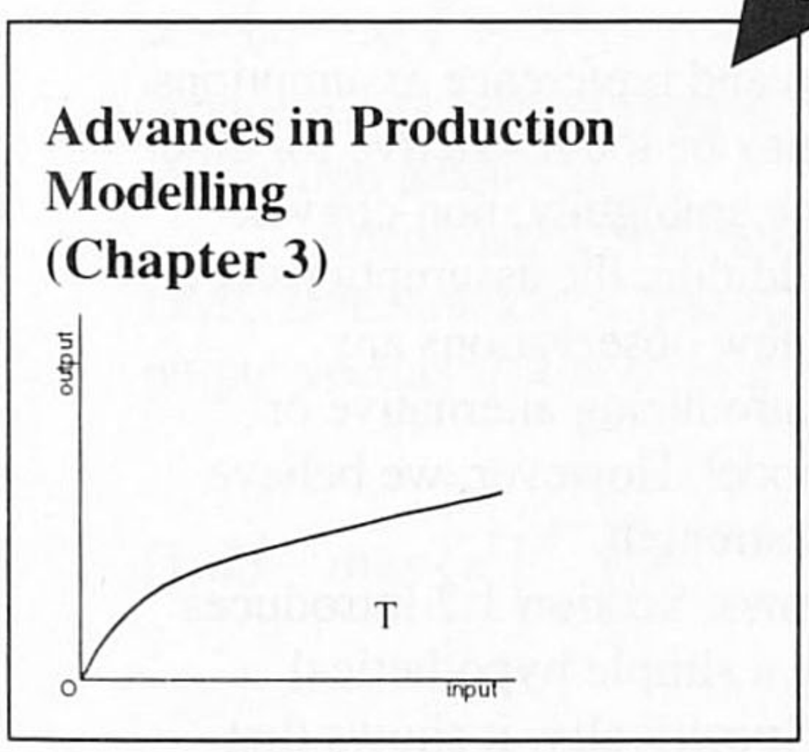
In addition, I thank Dr. Paul de Boer, Dr. Jan Brinkhuis, Prof. William Cooper, David Dekker, Prof. Rommert Dekker, Prof. Teun Kloek, Timo Kuosmanen (my non-convex brother-in-arms), Prof. Shlomo Maital, Prof. Philip Molyneux, prof. Leopold Simar, Prof. Emmanuel Thanassoulis, and Dr. Philip Thijsse for helpful comments and stimulating discussions, as well as various anonymous referees of *Operations Research* and *European Journal of Operational Research*.

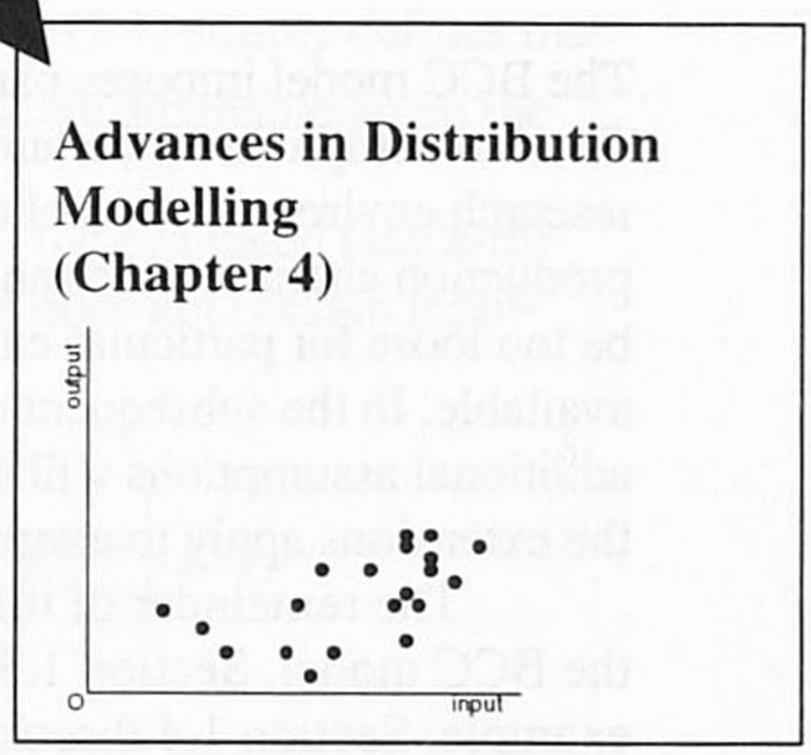
Last but not least, I thank Martine, who has greatly enriched my off-campus life in many ways. She even did some desperate attempts to understand my research. I strongly believe she deserves an honorary degree in Love, Patience and Understanding.

Chapter 2:
Advances in Preference Modelling









1.1. Introduction

This chapter provides the foundation for the chapters that follow. It discusses the basic concepts and characteristics of the DEA methodology, and in addition introduces the notation that will be used throughout the study. For non-experts, reading this chapter is a prerequisite for comprehending the developments discussed in the subsequent chapters. However, the latter chapters can be read independent of each other. This feature allows readers to take the most efficient route to specific subjects of interest.

We intend to demonstrate that DEA generates detailed information on the production technology and the efficiency of DMUs. That information includes efficiency measures, reference units and shadow prices. This information is tailored to each DMU individually, and in addition does not require a parametric specification of the production technology and the statistical distribution of the inefficiencies. These features give DEA a comparative advantage relative to Econometric Frontier Estimation, which focuses on obtaining a global representation of the production possibilities and in addition requires a specific functional form for the production technology and a specific distribution unction for the inefficiencies.

The DEA methodology comprises a wide variety of mathematical programming models for Efficient Frontier Analysis. In general, an empirical production possibilities set is constructed from a set of production data, and observed performance is compared with a particular efficient reference point taken from that empirical set. However, the models differ widely with respect to the assumptions imposed on the objectives of the DMUs, the production possibilities, and the statistical distribution of the observations. Notwithstanding this variety, we focus on the output-oriented version of the Banker, Charnes and Cooper (1984) model in this chapter. The BCC model has become the standard in the DEA methodology. One attractive feature of that specific model is that it relaxes the constant-returns-to-scale assumption imposed by its predecessor, the seminal Charnes, Cooper and Rhodes (1978) model, and allows for variable-returns-to-scale properties[12]. Many methodological chapters focus on extending the BCC model and many applications use it. In addition, the BCC model is embedded in most of the software packages currently available in the market[13].

The BCC model imposes particular production, distribution and preference assumptions. Those assumptions fit certain research environments, but may be too restrictive for other research environments, including those with high preference ambiguity, non-convex production characteristics and stochastic disturbances. In addition, the assumptions can be too loose for particular environments, especially if only few observations are available. In the subsequent chapters, the possibilities for introducing alternative or additional assumptions will be discussed in terms of this model. However, we believe the extensions apply to alternative DEA models with equal strength.

The remainder of this chapter is organized as follows. Section 1.2 introduces the BCC model. Section 1.3 illustrates its operation using a simple hypothetical example. Section 1.4 discusses its statistical properties. Specifically, it shows that

---

[12] Although BCC relaxes the constant returns-to-scale assumption, it does assume convexity, and hence excludes increasing returns-to-scale properties. Therefore, the BCC model cannot be considered a 'true' variable returns-to-scale model. For a more detailed discussion of violations of convexity and non-convex models, see Chapter 3.

[13] For software packages that are currently available, see Charnes *et al.* (1994) and Ali Emrouznejad's DEA homepage at http://www.csv.warwick.ac.uk/~bsrlu/.

DEA gives statistically consistent estimates for a wide range of production and distribution structures, and in addition discusses various approaches to analyzing the sampling distribution. Section 1.5 discusses the dual model formulation. DEA models have a Linear Programming structure that involves a primal and a dual formulation. The primal focuses on the selection of an efficient reference unit, whereas the dual focuses on the shadow prices at the efficient frontier. Section 1.6 discusses the use of slack corrections. Many DEA models include a radial objective. That objective does not guarantee the selection of a Pareto-Koopmans efficient reference unit. Slack corrections can remedy this problem. Finally, Section 1.7 offers some concluding remarks.

1.2. The standard model

DEA is predominantly used for analysing productive efficiency[14]. Most work in DEA adheres to the Pareto-Koopmans definition of technical efficiency. Koopmans (1951) provides a formal definition of this efficiency concept:

> *'A producer is technically efficient if an increase in any output requires a reduction in at least one other output or an increase in at least one input, and if a reduction in any input requires an increase in at least one other input or a reduction in at least one output.'*

Obviously, this definition can not be operationalised without specifying the production possibilities first. In this study, we represent the production possibilities by the following production possibilities set[15]:

$$\textbf{(1. 1)} \quad T = \left\{ (y,x) \in \Re_+^{p+m} \middle| \text{input } x \text{ can produce output } y \right\}.$$

In this expression, $y = \left(y_1 \cdots y_p\right)^T \in \Re_+^p$ represents an output vector and $x = \left(x_1 \cdots x_m\right)^T \in \Re_+^m$ represents an input vector. So, definition (1.1) simply defines the production possibilities as the set of input-output vectors that are attainable given the production technology.

According to the Pareto-Koopmans definition above, the evaluated (say *k*-th) DMU is technically efficient if and only if the production set does not contain input-output vectors that involve more output and less input, i.e.:

$$\textbf{(1. 2)} \quad \max_{x,y,s',s} \left\{ e^T \begin{pmatrix} s' \\ s \end{pmatrix} \middle| \begin{pmatrix} y - s' \\ -x - s \end{pmatrix} = \begin{pmatrix} y_k \\ -x_k \end{pmatrix}; (y,x) \in T; (s',s) \in \Re_+^{p+m} \right\} = 0.$$

---

[14] Obviously, the integral evaluation of organisational performance comprises more aspects than efficiency, including profitability, risk and quality. These dimensions can be included in DEA, and a number of studies have focussed on doing so. For example, Banker and Maindiratta (1988) included profitability, Olesen and Petersen (1995b) included quality and Post and Spronk (1999b) included risk. Chapter 2 discusses the inclusion of objectives in addition to technical efficiency in greater detail.

[15] Throughout the text, we use $\Re_+$ to denote the nonnegative orthant of the set of real numbers.

Here, $s' = (s'_1 \cdots s'_p)^T \in \Re^p_+$ and $s = (s_1 \cdots s_m)^T \in \Re^m_+$ represent slacks for the output and input variables respectively and $e = (1 \cdots 1)^T$ represents a unity vector of dimension $(p+m)\times 1$.

This definition can distinguish efficient from inefficient DMUs. However, it does not measure the magnitude of inefficiencies. A common efficiency measure is the Debreu (1951)-Farrell (1957) output efficiency measure, which measures the maximum equiproportionate output augmentation that is attainable given the production set without increasing any of the inputs. More formally, Debreu-Farrell output efficiency is defined as follows[16]:

$$\textbf{(1. 3)} \quad \theta_k = \max_{\theta} \{\theta | (\theta y_k, x_k) \in T\}.$$

A popular alternative measure is the Debreu-Farrell input efficiency measure, which measures the maximum equiproportionate input reduction that is attainable without reducing any of the outputs, i.e.:

$$\textbf{(1. 4)} \quad \rho_k = \min_{\rho} \{\rho | (y_k, \rho x_k) \in T\}.$$

Both measures are so-called 'radial' measures that focus on equiproportionate improvements in inputs and outputs. The original DEA models are formulated in terms of these measures. Radial measures are attractive, because they consider multiple inputs and outputs. In addition, radial measures are independent of the unit of measurement. However, radial measures do have an important disadvantage. Particular input-output vectors that are demonstrably inefficient in Pareto-Koopmans sense can be labelled efficient. Specifically, if individual inputs or outputs can be improved, but no equiproportionate improvement is feasible, a Pareto-Koopmans inefficient unit is Debreu-Farrell efficient. To circumvent this problem, a non-Archimedean slack correction can be included in the standard models. Section 1.5 deals with such corrections in more detail. In addition, a wide variety of non-radial measures have been proposed. For a discussion of alternative efficiency measures in DEA, see Boyd and Färe (1984), Cooper and Tone (1997), Färe, Grosskopf and Lovell (1985, 1994), Lovell and Pastor (1995). In addition, Färe and Lovell (1978), Bol (1986, 1988) and Russell (1985, 1988, 1990) give interesting discussions of desirable properties of efficiency measures.

In this chapter, the focus is on the radial output measure for simplicity. In addition, all extensions discussed and developed in this study are formulated in terms of a model that uses this measure. However, we believe the extensions apply to alternative efficiency measures with equal strength.

Unfortunately, one typically faces incomplete information about the production possibilities in empirical applications. Therefore, the efficiency measures can not be computed directly for a given input-output vector. To circumvent this problem, DEA uses particular empirical production sets to approximate the true production set. The empirical production sets are constructed using observations on a number (say n) of

---

[16] This output measure takes a value from $[1,\infty)$. Sometimes it is desirable to obtain a measure that takes values from $(0,1]$. Simply taking the inverse of θ_k suffices for such a measure. In fact, that operation gives the direct output distance function proposed by Shephard (1953, 1970).

comparable DMUs. Let the production data of the DMUs be represented by output matrix $Y = (y_1 \cdots y_n)$, with $y_j = (y_{1j} \cdots y_{pj})^T$, and input matrix $X = (x_1 \cdots x_n)$, with $x_j = (x_{1j} \cdots x_{mj})^T$. In addition, let D represent an index set of cardinality n.

In general, an empirical production set is constructed from the observations as the smallest subset in input-output space ($\Re_+^{p+m}$) that is consistent with particular assumptions about the production possibilities and the distribution of the observations in the data set.

The standard BCC model imposes four key assumptions:

(1) *homogeneity,*

(2) *free disposability,*

(3) *convexity*, and

(4) *accuracy.*

First, BCC assumes that the production set contains all observations in the data set, i.e.:

(1. 5) $(Y, X) \in T$.

This assumption corresponds to assuming that all DMUs operate under the same production technology. Therefore, we refer to this assumption as the *homogeneity* assumption.

Second, BCC assumes that it is possible, for each attainable input-output combination, to become less efficient by reducing output and/or increasing input. This so-called *free disposability* assumption, can be formalised as follows:

(1. 6) $(y,x) \in T \Rightarrow (y',x') \in T \ \forall (y',x') \in \Re_+^{p+m} : y' \leq y, x' \geq x$.

Third, the BCC model assumes that the production set is *convex*, i.e.:

(1. 7) $(Y, X) \subset T \Rightarrow (Y\lambda, X\lambda) \in T \ \forall \lambda \in \Re_+^n : e^T \lambda = 1$.

Finally, there is an implicit distribution assumption in the standard model. Recall that the output matrix Y and the input matrix X represent the production data. This is not a harmless issue of notation. Equating the data set with the true values of the inputs and outputs effectively assumes that the input-output values are *accurately measured*, and excludes the effects of measurement error. Chapter 4 relaxes this assumption and assumes that estimates rather than the true values are available. More formally, using $(\hat{Y}, \hat{X})$ to denote the estimated input-output matrices, the accuracy assumption amounts to:

(1. 8) $(\hat{Y}, \hat{X}) = (Y, X)$.

Models that impose additional or alternative production and distribution assumptions are discussed in Section 3 and 4 respectively. Especially the assumptions of homogeneity, convexity and accuracy have been subject of much discussion.

The BCC model uses the smallest subset in input-output space that is consistent with the above assumptions as an empirical production set to approximate the true production set. More formally, BCC uses the following set as an approximation:

$$\textbf{(1. 9)} \quad \hat{T}_{BCC} = \left\{ (y,x) \in \Re_+^{p+m} \middle| \begin{pmatrix} y \\ -x \end{pmatrix} \leq \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e = 1; \lambda \in \Re_+^n \right\}.$$

Measuring technical efficiency relative to this set gives the following estimator:

$$\textbf{(1. 10)} \quad \hat{\theta}_{BCC,k} = \max_{\theta} \left\{ \theta \middle| (\theta y_k, x_k) \in \hat{T}_{BCC} \right\}$$

$$= \max_{\lambda,\theta} \left\{ \theta \middle| \begin{pmatrix} \theta y_k \\ -x_k \end{pmatrix} \leq \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e = 1; \lambda \in \Re_+^n \right\}.$$

Solving the following Linear Programming problem can obtain this estimator:

$$\textbf{(1. 11)} \qquad \hat{\theta}_{BCC,k} = \max_{\lambda,\theta,s,s'} \theta$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} - s_r = \theta y_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i = x_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s_r' \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

This model selects the convex combination of observations that maximises the radial output augmentation relative to the evaluated unit, while consuming an amount of input that is less or equal to that of the evaluated unit.

The optimal θ is a summary measure for efficiency. In addition, the optimal slacks s', s represent non-radial inefficiencies. The efficiency measures can be used as a summary measure for the efficiency of past operations and for assessing potential improvements for future operations. Hence, they provide valuable information for managerial planning and control. In addition, academic researchers can use the efficiency scores to test economic theories.

The optimal convex combination, i.e. λ^*, can be interpreted as a reference unit that represents best practice. The reference unit can be an observed DMU or a

fictitious DMU, constructed as a combinations of observed DMUs[17,18]. Comparing the work process of the evaluated unit with that of this reference unit can reveal causes and remedies for inefficiencies. By identifying reference units in addition to measuring inefficiency, DEA has a comparative advantage relative to Econometric Frontier Estimation.

In addition to efficiency measures and reference units, DEA models can be used to compute shadow prices or virtual weights for the inputs and outputs. Specifically, the dual model formulation discussed in Chapter 1.6 focuses on this information. The shadow prices are computed as the marginal production possibilities (for the inputs) and the marginal input requirements (for outputs) at the reference point on the empirical frontier. The shadow prices can be used for measuring local scale, substitution and transformation characteristics of the production technology, and hence provide valuable information on the production technology for production planning and economic analysis.

In contrast to parametric approaches whose objective it is to optimise a single regression plane that is assumed to apply to each DMU, DEA optimises the efficiency measure of each individual DMU. For each DMU in the data set, DEA solves a separate mathematical programming problem. Therefore, the information obtained is specially tailored to each individual DMU.

1.3. Example illustration

To illustrate the operation of DEA, consider 5 hypothetical DMUs that operate under a single-input single-output technology and with input-output data displayed in Table 1.1.

	X	Y
DMU A	4	4
DMU B	8	4
DMU C	10	8
DMU D	12	7
DMU E	18	9

Table 1. 1 A hypothetical data set

Graph 1.1 displays these data relative to the hypothetical production set *T*. The solid line represents the assumed production frontier.

---

[17] The production assumptions determine what kind of fictitious DMUs DEA models use. For example, the BCC model allows for convex combinations, and the CCR model allows for positive linear combinations of observations. Chapter 3 discusses this issue in more detail.

[18] If the production assumptions hold, the fictitious DMUs can be demonstrated to represent attainable input-output levels. Still, fictitious DMUs may be less acceptable and defensible as performance benchmarks than observed DMUs. This consideration is important in managerial applications, and gives an argument for using the FDH model (see Chapter 3), which excludes fictitious DMUs and considers observed DMUs only. Nevertheless, for the purpose of efficiency estimation the consideration is less important, and fictitious DMUs can be used without harm.

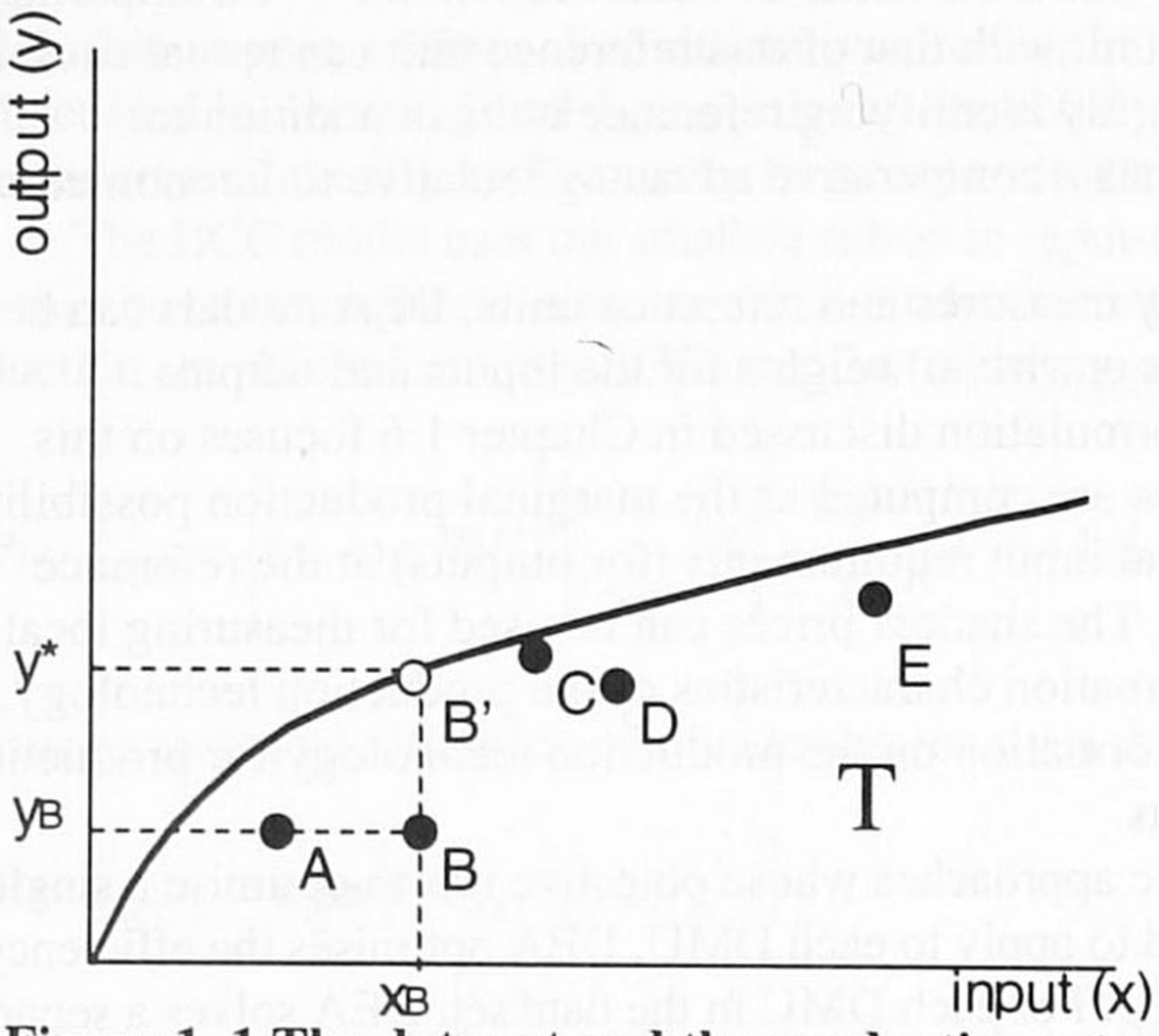


Figure 1. 1 The data set and the production set.

Unit B clearly is technically inefficient, as its output (y_B) falls substantially short of the maximum output attainable at its input (x_B), y^*, giving an efficiency measure of $\theta_B = y^* / y_B$. Figure 1.2 displays the BCC empirical production set (T_{BCC}). The piecewise linear line represents the estimated frontier derived by solving the DEA model for various input values.

For example, solving it for DMU B gives the fictitious DMU B', which is constructed as an equally weighed average of DMU A and C. This unit consumes the same amount of input as B, but produces substantially more output. Comparing B's output with that of its reference unit gives an efficiency estimate of $\hat{\theta}_B = y_{B'} / y_B$.

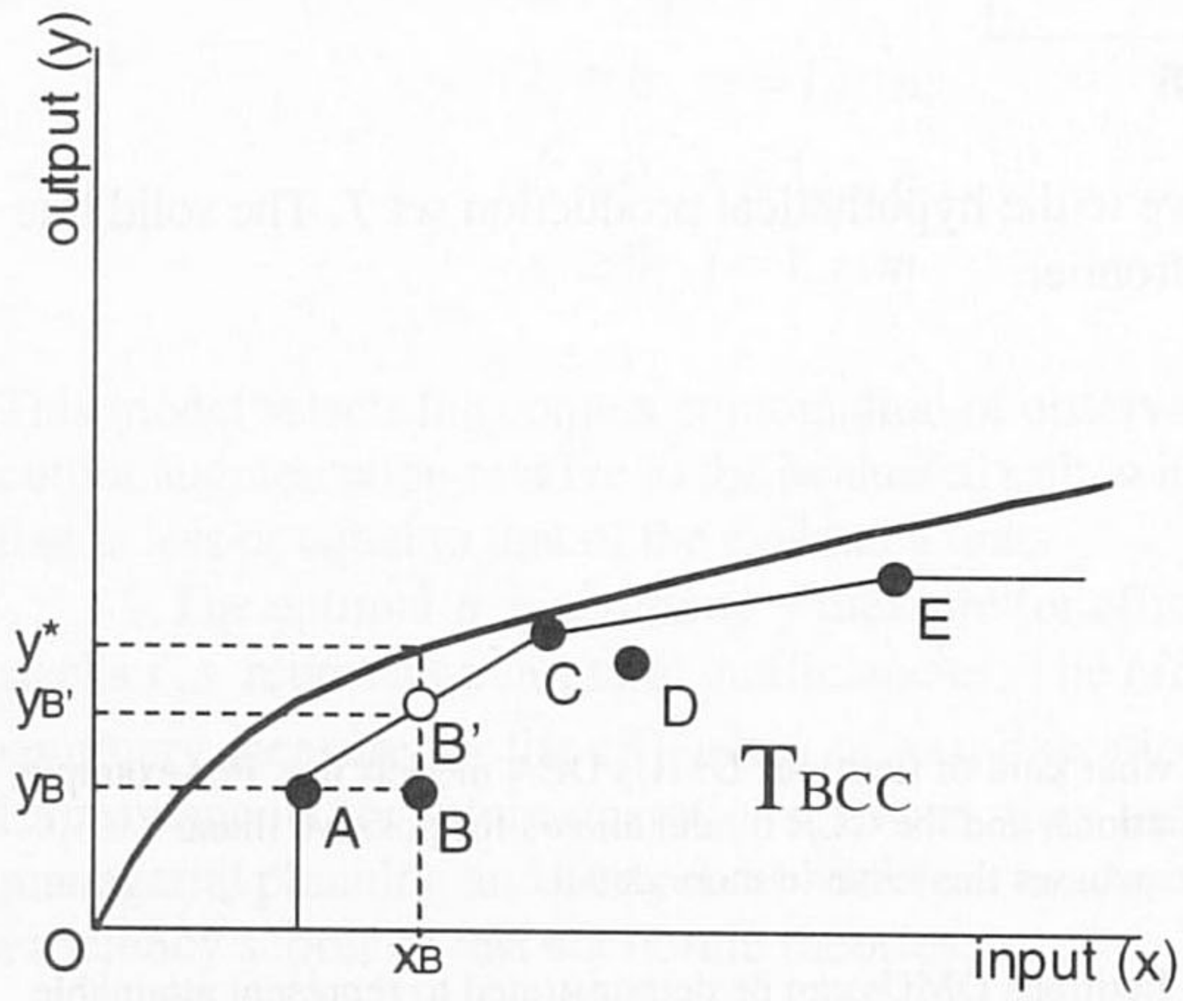


Figure 1. 2 The BCC model for the example data set

1.4. Statistical properties

Much research effort has been directed at analysing the statistical properties of DEA efficiency estimates. As discussed above, standard DEA models assume accurate measurement for the input-output vector, and therefore are deterministic by nature. Nevertheless, these models do have a statistical basis. After all, efficiency is measured relative to an empirical estimate of the true production set. If the DMUs in the data set are considered as a random sample from a hypothetical infinite universe, computation is subject to sampling error[19]. The effect of sampling error has been studied extensively. Grosskopf (1996), Simar (1996) and Simar and Wilson (1999a) provide a comprehensive survey of statistical inference in DEA. This section discusses some of the key results, but it is not intended to be thorough discussion of the subject (which merits a study in its own right). Rather, the objective is to provide the basic statistical knowledge to comprehend the discussion in the subsequent chapters. The interested reader can find more details in the references provided.

One interesting result is that the DEA estimators can be demonstrated to be *statistically consistent* (i.e., roughly speaking, asymptotically unbiased and with a vanishing variance) for a wide range of production and distribution structures. The key to understanding this result is that the empirical production set is contained within the true production set, i.e. $\hat{T}_{BCC} \subseteq T$, if the standard production assumptions (free disposability (1.6) and convexity (1.7)) and distribution assumptions (homogeneity (1.5) and accuracy (1.8)) are being satisfied. Therefore, the efficiency estimator is biased below true efficiency, i.e.[20]:

$$\textbf{(1. 12)} \qquad P\left(\hat{\theta}_{BCC,k} - \theta_k > \varepsilon\right) = 0 \quad \forall \varepsilon > 0 .$$

Consequently, $\hat{\theta}_{BCC,k} = 1$ is a necessary condition for efficiency. In addition, increasing the number of observations in the data set generally improves the estimate. This is because the empirical production set associated with a larger data set generally represents a larger subset of the true production set (provided the standard assumptions are being satisfied). In addition, a larger subset of the production set is associated with a better efficiency estimate, i.e.:

$$\textbf{(1. 13)} \qquad \hat{T}_{BCC} \subseteq \hat{T}^*_{BCC} \subseteq T \Rightarrow \left|\hat{\theta}^*_{BCC,k} - \theta_k\right| \leq \left|\hat{\theta}_{BCC,k} - \theta_k\right|.$$

---

[19] Measurement error constitutes a source of error in addition to sampling error. It is important to distinguish sampling error from measurement error, as the two sources of error have different effects on the goodness of the estimators. Sampling error arises if the input-output vectors in the sample do not give a good representation of the production possibilities. Increasing the sample size generally can remedy this problem. By contrast, measurement error arises if the observed values of the input-output vectors do not represent the actual achievements. In general, increasing the sample size can't remedy this problem. Chapter 4 discusses the problems associated with measurement error in more detail, and discusses stochastic DEA models that do account for measurement error in addition to sampling error.

[20] Note that efficiency is measured here as the maximum equiproportionate output augmentation, i.e. Debreu-Farrell output efficiency (1.3). The higher this measure the less efficient a DMU is. Therefore 'overestimation of efficiency' refers to the DMU appearing less efficient than it actually is. By contrast, if efficiency is measured as the direct output distance measure by Shephard (1953, 1970), i.e. the inverse of Debreu-Farrell output efficiency, a higher value is associated with more efficiency.

For example, including additional DMUs in the example data set discussed in Section 1.3 could give the situation displayed in Figure 1.3. The empirical production set is substantially closer to the true production set than that in Figure 1.2. In addition, the reference unit B' is closer to the frontier and the efficiency estimate for DMU B improves.

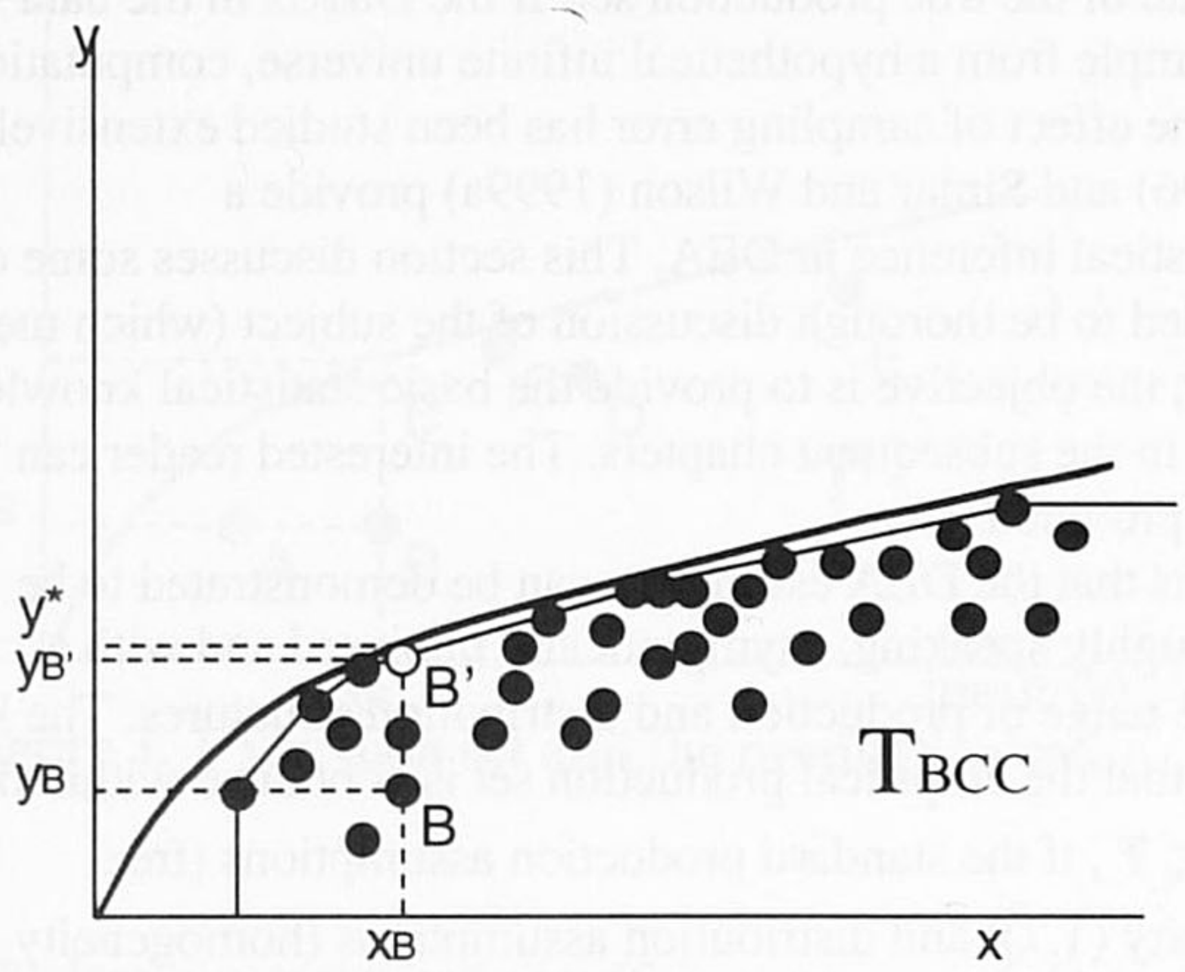


Figure 1. 3 The effect of increasing the sample size

Minimal assumptions can demonstrate that the efficiency estimator in fact converges to true efficiency as the sample size goes to infinity. More specifically, the efficiency estimator is statistically consistent for a wide range of distributions for the observations within the production set, i.e.:

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_{BCC,k} - \theta_k\right| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0. \tag{1. 14}$$

Banker (1993) and Korostelev, Simar and Tsybakov (1995a, 1995b) discuss consistency for the single-output case, and Kneip, Park and Simar (1998) discuss it for the general multi-output case.

For deriving the statistical properties of the estimators, this study considers the observations as identically and independently distributed random variables with a strictly positive probability of dominating arbitrary input-output combinations from the interior of the production set, i.e.[21]:

$$P(y_j \geq y, x_j \leq x) = F(y,x) > 0 \quad \forall (y,x) \in T; j \in D. \tag{1. 15}$$

This condition is sufficient for consistency.

---

[21] Note that F does *not* denote the cumulative density function $P(y_j \leq y, x_j \leq x)$, but rather $P(y_j \geq y, x_j \leq x)$.

Proof 2.1 If the data set contains a DMU that consumes at most x_k to produce at least $\theta_k y_k$, the estimator necessarily exceeds true efficiency, i.e.:

(I) $\exists j \in D : x_j \leq x_k ; y_j \geq \theta_k y_k \Rightarrow \hat{\theta}_{BCC,k} \geq \theta_k .$

In addition, $(\theta_k y_k, x_k) \in T$ holds by definition. Therefore, (1.15) implies $P(y_j \geq \theta_k y_k, x_j \leq x_k) = F(\theta_k y_k, x) > 0$. Consequently, the probability of observing at least one such DMU converges to unity as the sample increases, i.e.:

$$\lim_{n\to\infty} P\left(\exists j \in D : x_j \leq x_k ; y_j \geq \theta_k y_k\right) = \lim_{n\to\infty} \left(1 - (1 - F(x_k, \theta_k y_k))^n\right) = 1 .$$

Combining this with (I), we find that, asymptotically, the estimator is biased above true efficiency:

(II) $$\lim_{n\to\infty} P\left(\hat{\theta}_{BCC,k} - \theta_k < \varepsilon\right) = 0 \quad \forall \varepsilon > 0.$$

Combining (II) and (1.12) gives consistency, i.e.:

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_{BCC,k} - \theta_k\right| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0. \square$$

Although the DEA estimator can be demonstrated to be consistent for general production and distribution structures, it can suffer from substantial *finite sample error*. In finite samples, best practice generally falls short of potential production for at least some of the observations, and consequently efficiency is underestimated for those observations. This is especially true if the number of observations is relatively small relative to the number of input and output variables (*'the curse of dimensionality'*). The positive effect of sample size and the negative effect of dimensionality are also reflected in the *asymptotic rate of convergence* of the estimator, which equals $n^{-\left(\frac{2}{p+m+1}\right)}$. Kneip *et al.* (1998) and Gijbels, Mammen, Park and Simar (1999) provide a comprehensive discussion of this subject.

To account for finite sample error, it is desirable to analyse *the sample distribution* of the estimators. Knowledge of the sampling distribution allows for bias correction and the estimation of confidence intervals for the efficiency estimates. Two approaches have been proposed for analysing the sampling distribution. First, Gijbels, Mammen, Park and Simar (1999) derived the *asymptotic sampling distribution* for the estimators. However, they provide results for the single-input single-output case only, and it is not clear how to generalise their results to the general multi-input multi-output case. In addition, their approach requires the estimation of unknown distribution parameters, which introduces additional statistical noise and imprecision. Second, it has been proposed to approximate the sampling distribution by means of *bootstrapping*. Bootstrapping involves the repeated simulation of the data generating

process and the application of the original estimator to each simulated sample so that the resulting estimators mimic the sampling distribution of the original estimator. Atkinson and Wilson (1995), Gstach (1995), Lötgren (1997), Hall, Härdle and Simar (1995), Ferrier and Hirschberg (1997), and Simar and Wilson (1998a, 1998b, 1999b, 1999c, 1999d) provide a comprehensive discussion of this subject.

1.5. Slack correction

Model (1.10) uses the Debreu-Farrell output efficiency measure. Unfortunately, maximising the radial output augmentation does not necessarily yield a Pareto-Koopmans efficient reference unit. More specifically, multiple optimal solutions may exist, i.e. multiple reference units that give an identical radial output augmentation, i.e. elements of the following set:

$$\textbf{(1. 16)}\quad \left\{\lambda \in \Re_+^n \middle| \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix} \geq \begin{pmatrix} \theta^* y_k \\ -x_k \end{pmatrix}; \lambda^T e = 1\right\}.$$

Figure 1.4 illustrates this point. The BCC model could compare DMU *F* with reference point *F'*, which is constructed as a weighed average of DMU *E* and DMU *G*. However, all convex combinations of DMU *E* and DMU *G* give the same optimal solution and therefore are potential reference units.

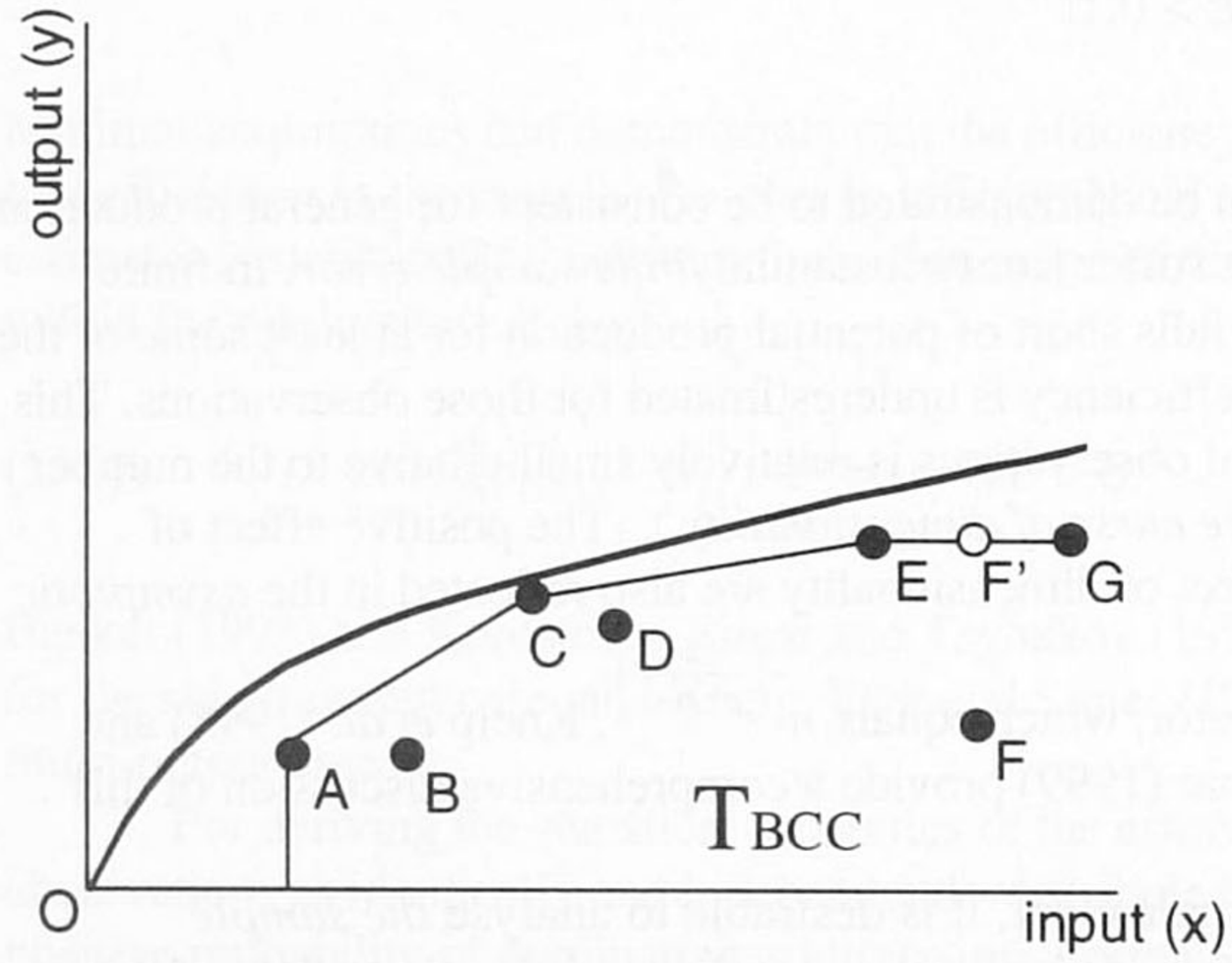


Figure 1. 4 The selection of a reference unit that is Debreu-Farrell efficient, but Pareto-Koopmans inefficient

Some of the alternative solutions may be dominated in Pareto-Koopmans sense, i.e. alternative solutions may have more output and/or less input. For example. in Figure 1.4, DMU *E* Pareto-Koopmans dominates all other convex combinations of DMU *E* and DMU *G*. However, only Pareto-Koopmans efficient reference units have to be

considered as potential reference units[22]. We will refer to these units as the BCC admissible set. More formally, that set is defined as:

$$\textbf{(1.17)}\quad A_{BCC}=\left\{\lambda\in\Re_+^n \middle| \left\{\tau\in\Re_+^n \middle| \begin{pmatrix} Y\lambda-s' \\ -X\lambda-s \end{pmatrix}=\begin{pmatrix} Y\tau \\ -X\tau \end{pmatrix}; \tau^T e=1; (s' \quad s)e>0; (s',s)\in\Re_+^{p+m}\right\}=\varnothing, \lambda^T e=1\right\}.$$

In this expression, $\tau=(1\cdots1)$represents an (n x 1) unity vector, and e' and e represent (p x 1) and (m x 1) unity vectors respectively. The optimal reference unit (λ^*) is the solution to the following problem:

$$\textbf{(1. 18)}\quad \max_{\lambda,\theta}\left\{\theta \middle| \begin{pmatrix} \theta y_k \\ -x_k \end{pmatrix}\leq\begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda\in A_{BCC}\right\}.$$

To solve this problem, it has been proposed to use the following slack corrected version of model (1.10):

$$\textbf{(1. 19)}\quad \theta_{BCC,k}=\max_{\theta}\left\{\theta+\varepsilon e^T\begin{pmatrix} s' \\ s \end{pmatrix} \middle| (\theta y_k+s', x_k-s)\in\hat{T}_{BCC}\right\}$$

$$=\max_{\theta,\lambda,s',s}\left\{\theta+\varepsilon e^T\begin{pmatrix} s' \\ s \end{pmatrix} \middle| \begin{pmatrix} \theta y-s' \\ -x_k-s \end{pmatrix}=\begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e=1; \lambda\in\Re_+^n; (s',s)\in\Re_+^{p+m}\right\}.$$

In this expression, ε represents a non-Archimedean infinitesimal small value. The slack correction does not interfere with the inefficiency estimate, but it pre-empts the selection of reference units that are dominated in Pareto-Koopmans sense by other reference units.

The slack-corrected model can be solved using the following Linear Programming problem:

---

[22] The rationale behind the use of Pareto-Koopmans efficiency is that the decision-maker can be assumed non-satiable in most research environments, i.e. the decision-maker prefers more output over less, and in addition prefers less input over more. Pareto-Koopmans inefficient reference units can increase outputs (or decrease inputs) without increasing inputs (decreasing outputs), and can therefore be excluded as decision alternatives for non-satiable decision-makers.

$$\textbf{(1. 20)}\quad \hat{\theta}_{BCC,k} = \max_{\lambda,\theta,s,s'} \theta + \varepsilon\left(\sum_{i=1}^{m} s_i + \sum_{r=1}^{p} s'_r\right)$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} - s_r = \theta y_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i = x_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s'_r \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

The slack-corrected model requires the selection of an appropriately small value for ε. To circumvent the associated ambiguity, the model can alternatively be solved by first maximising the radial gain and subsequently maximising a target sum to obtain further gains where possible, retaining all the radial gain[23]. Most DEA-software embeds such a two-step procedure. More specifically, the first stage is finding the optimal value for θ by solving the following problem:

$$\textbf{(1. 21)}\quad \theta^*_k = \max_{\lambda,\theta,s,s'} \theta$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} - s_r = \theta y_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i = x_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s'_r \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

---

[23] For computational aspects of DEA models, see Adolphson, Cornia and Walters (1990), Ali (1990, 1993, 1994), Ali and Seiford (1990, 1993), Charnes, *et al.* (1994), Färe and Hunsaker (1986), Sueyoshi (1990), and Sueyoshi and Chang (1989).

The computations for the applications and simulations in this study are performed using computer programs in GAUSS. Appendix 3.1 presents the computer program for the simulation strudy in Section 3.7. That program contains the standard BCC model in addition to the Free DisposableHull model (Deprins, Simar and Tulkens (1984)) and the Transconvex Data Envelopment Analysis model (Post (1997b, 1998), and Dekker and Post (1999))

The second stage uses the optimal solution, to find the optimal reference unit by solving the following problem:

$$\textbf{(1. 22)} \qquad \max_{\theta,\lambda_j,s_i,s_r'} \left(\sum_{r=1}^{p} s_r' + \sum_{i=1}^{m} s_i \right)$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} - s_r' = \theta_k^* y_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i = x_{ik} \quad i = 1,\ldots,m;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s_r' \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

If the data set does not contain multiple reference units with the optimal input-output values, the variance correction does not have an effect on the optimal solution. However, if the data set does contain a multitude of such reference units, the slack correction ensures Pareto-Koopmans efficiency.

1.6. Dual formulation

The models discussed so far have focussed on measuring efficiency and identifying an efficient reference unit. The dual formulation of DEA models provides additional information that can be used for evaluating local scale, substitution and transformation characteristics of the production set.

The primal model formulation measures efficiency as the maximum distance from a convex combination of observations. By contrast, the dual model measures efficiency as the minimum distance from a linear hyperplane that supports the observations from above. More specifically, the dual selects a linear hyperplane H with intercept $u \in \Re^1$, positive output weights $\mu \in \Re_+^p$ and positive input weights $v \in \Re_+^m$, i.e.:

$$\textbf{(1. 23)} \quad H = \left\{ (y,x) \in \Re_+^{p+m} \middle| \mu^T Y - v^T X = u; (\mu,v) \in \Re_+^{p+m}; u \in \Re^1 \right\}$$

That hyperplane is required to support all observations from above, i.e.:

$$\textbf{(1. 24)} \quad (Y,X) \in \left\{ (y,x) \in \Re_+^{p+m} \middle| \mu^T Y - v^T X \leq u; (\mu,v) \in \Re_+^{p+m}; u \in \Re^1 \right\} \quad \forall j \in D.$$

Efficiency is defined as the minimum distance from such a hyperplane that supports the observations from above, i.e.:

$$\textbf{(1. 25)} \quad \hat{\theta}_{BCC,k}^D = \min_{u,v,\mu} \left\{ u + v^T x_k \middle| \mu^T y_k = 1; \mu^T Y - v^T X \leq u; (\mu,v) \in \Re_+^{p+m} \right\}$$

Note that the weights are restricted by $\mu^T y_k = 1$. This restriction standardises the

efficiency score, so that it takes a value from the interval $[1,\infty)$.

This measure can be computed by solving the following LP problem:

$$\textbf{(1. 26)}\quad \hat{\theta}^{D}_{BCC,k} = \min_{u,v,w} u + \sum_{i=1}^{m} v_i x_{ik}$$

$$\text{s.t.} \sum_{r=1}^{q} \mu_r y_{rk} = 1;$$

$$u + \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{q} \mu_r y_{rj} \le 0 \quad j = 1,\ldots,n;$$

$$v_r, \mu_r \ge 0 \quad i = 1,\ldots,m; r = 1,\ldots,q.$$

Figure 1.5 illustrates the operation of the dual model formulation. The dual minimises efficiency of B by selecting the positively sloped linear line that supports all data, that is closest to B. This line is represented by $y = u^* + v^* x$, where u^* and v^* represent optimal values.

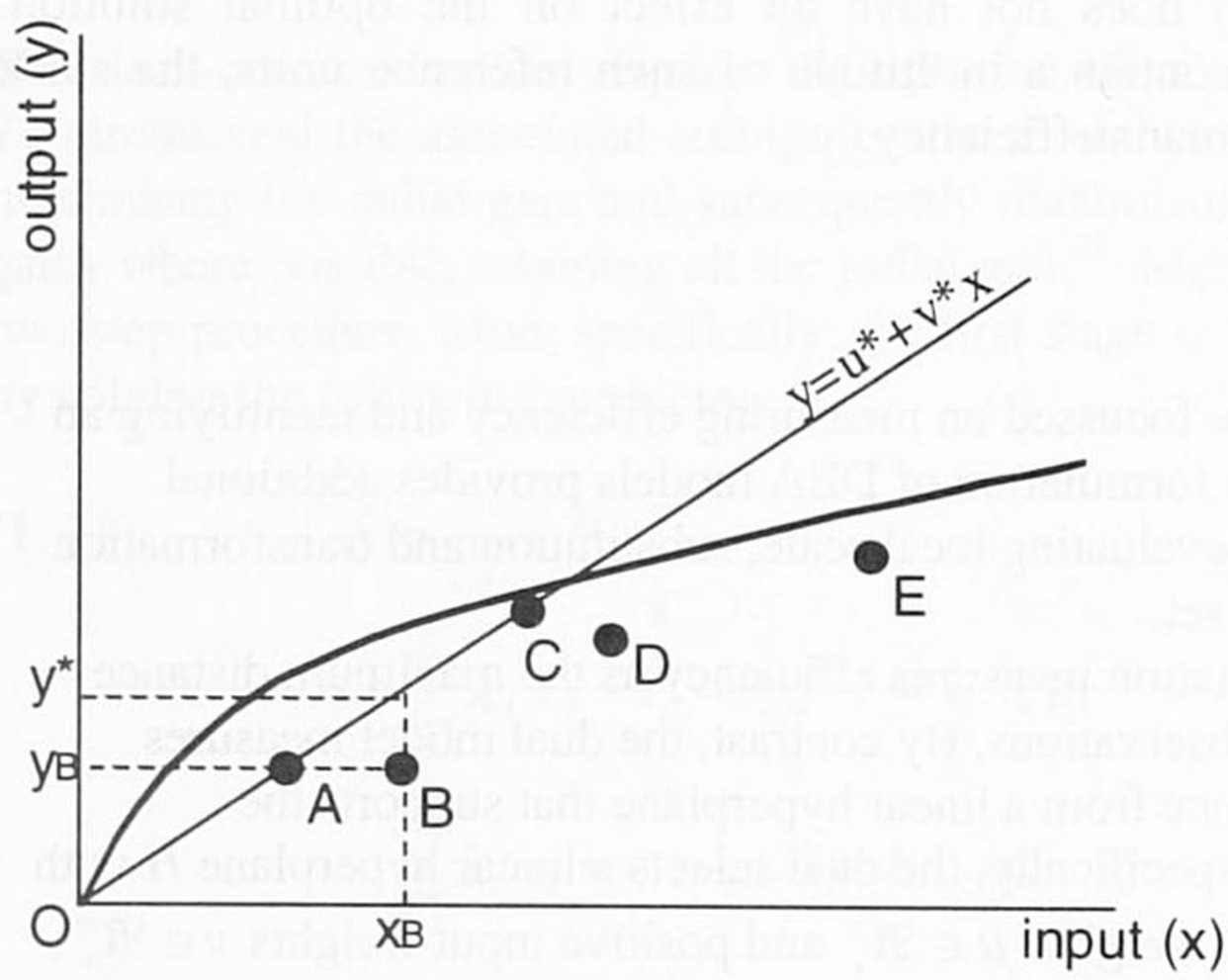


Figure 1. 5 The operation of the dual formulation of the BCC model.

Since the primal and dual models have Linear Programming structures, and, in addition, are formulated such that feasible solutions exist[24], the duality theorem from Linear Programming theory implies that dual and primal solutions are equal. For example, the output distance of DMU B from the optimal hyperplane in Figure 1.5 is identical to the output distance from the optimal convex combination in Figure 1.4.

However, the dual does provide additional information. The optimal weights can be used to estimate local scale and substitution characteristics of the frontier (e.g. Lewin and Morey (1981), Banker, Conrad and Strauss (1986), and Charnes, Cooper, Huang and Sun (1990)). As the supporting hyperplane with $u = 0$ is a ray emanating from the

[24] The evaluated DMU is included in the data set used to construct the empirical production set. Therefore, it is possible to compare the evaluated DMU with itself. That solution necessarily satisfies all restrictions imposed on reference units.

origin, inputs and outputs can freely be scaled without violating the set. This indicates the local existence of constant returns-to-scale. By contrast, $u < 0$ implies that inputs and outputs can be scaled up and therefore indicates local non-decreasing returns-to-scale. In similar vein, $u < 0$ indicates local non-increasing returns-to-scale[25]. Similarly, the input-output weights v and μ can be used to estimate substitution rates.

However, there are some complications associated with this approach. For example, there may exist multiple optimal solutions, and in addition lack of variation in the data can lead to hyperplanes that are not of full dimension. For approaches that guarantee a set of well-defined local scale and substitution characteristics, see Olesen and Petersen (1996), Green, Doyle and Cook (1996), Seiford and Zhu (1998), and Räty (1998).

We focus on the primal formulation in this study. This is not a restriction, because the optimal values of the dual variables can be derived directly from the primal variables.

1.7. Concluding remarks

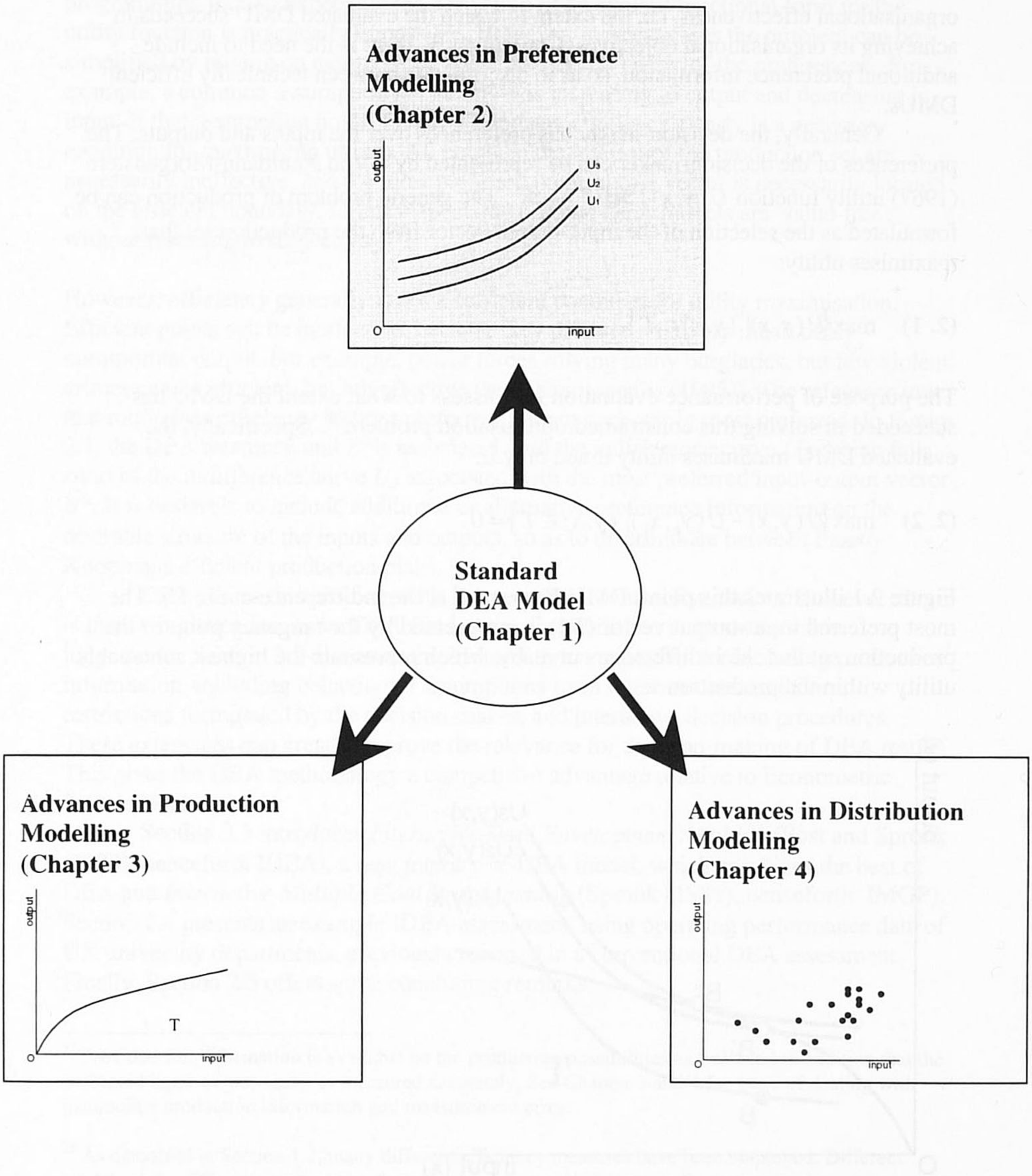
This chapter discussed the basic concepts and characteristics of the DEA methodology. In addition, it introduced the notation that will be used throughout the study. We have demonstrated that DEA constructs an empirical production set from a sample of comparable DMUs, and measures the efficiency of the DMUs relative to that set. We have demonstrated that DEA generates detailed information on the production technology and the efficiency of DMUs relative to that technology. This information includes efficiency measures, reference units and, for the dual model formulation, shadow prices. In contrast to econometric frontier estimation, this information is tailored to each DMU individually, and in addition does not require a parametric specification of the production technology and the distribution of the inefficiencies. Still, the goodness of that information critically depends on the goodness of the assumptions imposed with respect to the organizational objectives, the production technology, and the distribution of the observations. Assumptions are always required, to guarantee goal congruence, and in addition to reduce finite sample error.

For simplicity, we focused on the output-oriented version of the standard Banker, Charnes and Cooper (1984) model in this chapter. The BCC model imposes particular production, distribution and preference assumptions. Those assumptions fit certain problems, but may be too restrictive for other problems, including those with high preference ambiguity, non-convex production characteristics and stochastic disturbances. In the subsequent chapters, the possibilities for introducing alternative or additional assumptions will be discussed in terms of this model. Chapter 2 deals with preference assumptions, Chapter 3 with production assumptions, and Chapter 4 with distribution assumptions. This chapter provides the necessary foundation for the non-expert to comprehend these Chapters. However, these Chapters can be read independently of each other. This feature allows readers to take the most efficient route to specific subjects.

---

[25] For a more detailed discussion of measuring returns-to-scale properties using DEA, see Banker and Thrall (1992), Zhu and Shen (1995) and Banker *et al.* (1996).

Chapter 2:
Advances in Preference Modelling



2.1. Introduction

Most DEA models, including the standard BCC model discussed in the introduction, focus on technical efficiency. Technical efficiency generally is an important organisational objective. However, it generally is not the sole objective. Therefore, technical efficiency generally is a necessary, but not sufficient, condition for organisational effectiveness, i.e. the extent to which the evaluated DMU succeeds in achieving its organisational objectives. Consequently, there is the need to include additional preference information, so as to discriminate between technically efficient DMUs.

Generally, the decision-maker has preferences over the inputs and outputs. The preferences of the decision-maker can be represented by a Von Neumann-Morgenstern (1967) utility function $U(y,x):\Re_{+}^{p+m}\rightarrow\Re^{1}$. The general problem of production can be formulated as the selection of the input-output vector from the production set that maximises utility:

$$\textbf{(2. 1)}\quad \max_{x,y}\left\{U(y,x)\middle|\ (y,x)\in T\right\}.$$

The purpose of performance evaluation is to assess to what extent the DMU has succeeded in solving this constrained optimisation problem[26]. Specifically, the evaluated DMU maximises utility if and only if:

$$\textbf{(2. 2)}\quad \max_{x,y}\left\{U(y,x)-U(y_k,x_k)\middle|\ (y,x)\in T\right\}=0.$$

Figure 2.1 illustrates this point. DMU B operates at the indifference curve U_1. The most preferred input-output vector (B^*) is represented by the tangency point of the production set and the indifference curve U_3, which represents the highest achievable utility within the production set.

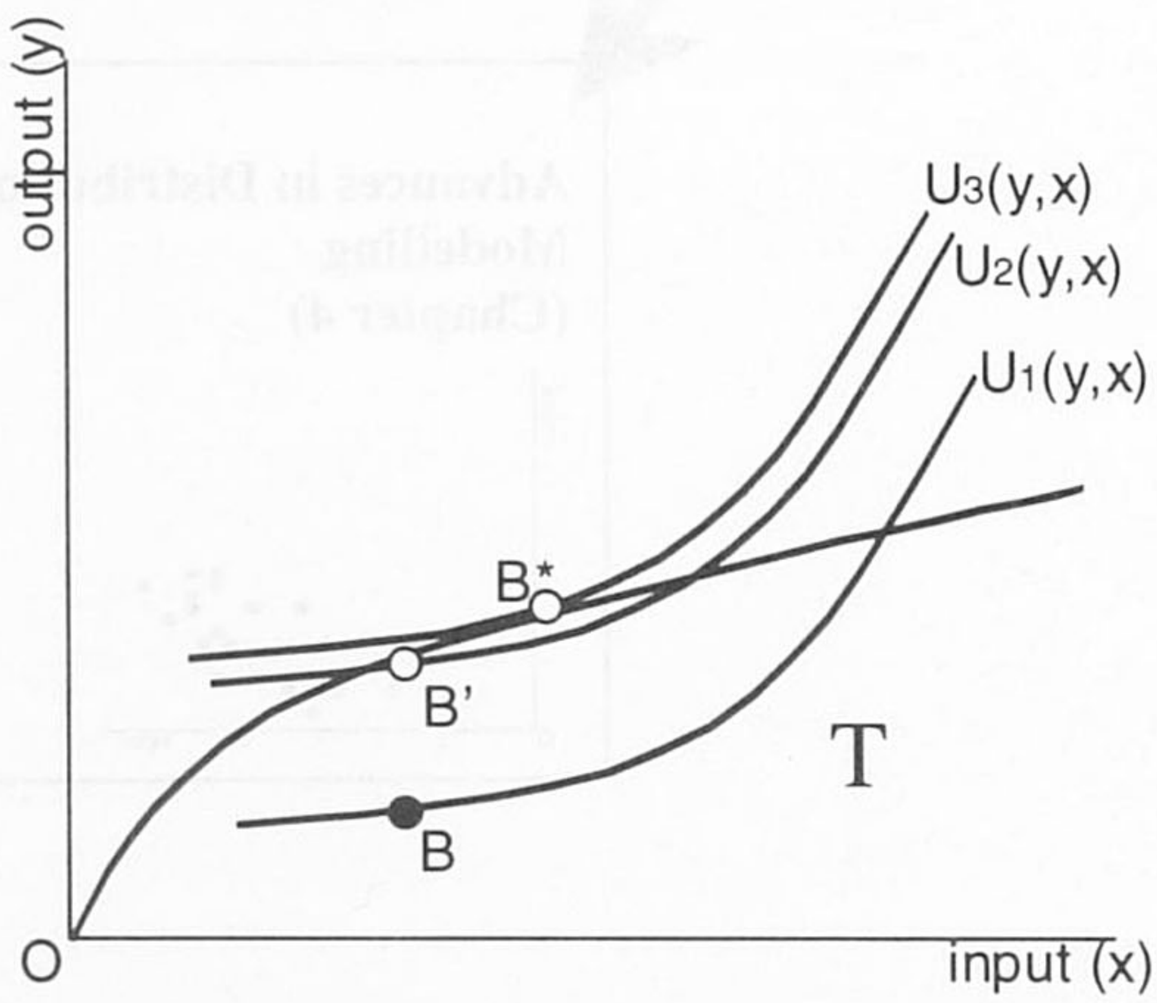


Figure 2. 1 Measuring effectiveness rather than technical efficiency

---

[26] As such, DEA forms part for a broader body of research dealing with the nonparametric testing of the consistency of observed economic behaviour with particular constrained optimisation problems. See footnote 6 in the introduction.

If the utility function is fully known, the performance evaluation problem could be solved using standard numerical techniques[27]. Unfortunately, preference information generally is incomplete. It is generally difficult to reliably elicit decision-maker preferences and in addition incorporate those preferences in a mathematical programming model. Especially the derivation of an explicit functional form for the utility function is practically impossible. However, in many cases the problem can be simplified by relying on assumptions about the general nature of the preferences. For example, a common assumption is that utility is increasing in output and decreasing in input. If that assumption holds, Pareto-Koopmans efficiency clearly is a necessary condition for maximising utility. All points in the interior of the production set are necessarily ineffective, and the most preferred input-output vector is necessarily located on the efficient boundary. In this respect, the original DEA models are 'value-free', without resorting to subjectivity [28].

However, efficiency generally is not a sufficient condition for utility maximisation. Efficient points can be ineffective, because they produce relatively much of an unimportant output. For example, police forces solving many burglaries, but few violent crimes can be efficient, but not effective (see Thannassoulis (1995)). The reference unit that minimises efficiency is more preferred, but not necessarily most preferred. In Figure 2.1, the DEA reference unit B' is associated with the indifference curve U_2, which falls short of the indifference curve U_3 associated with the most preferred input-output vector B^*. It is desiarble to include additional or alternative preference information on the desirable structure of the inputs and outputs, so as to discriminate between Pareto-Koopmans efficient production plans.

DEA shares this problem with Econometric Frontier Estimation. However, there is increasing attention in the DEA field to the desirability to incorporate value judgements. Section 2.2 discusses three alternative approaches for including preference information, including behavioural assumptions from micro-economic theory, restrictions formulated by the decision-maker, and interactive decision procedures. These extensions can greatly improve the relevance for decision-making of DEA results. This gives the DEA methodology a competitive advantage relative to Econometric Frontier Estimation.

Section 2.3 introduces *Interactive Data Envelopment Analysis* (Post and Spronk (1999), henceforth IDEA), a new interactive DEA model, which combines the best of DEA and *Interactive Multiple Goal Programming* (Spronk (1981), henceforth IMGP). Section 2.4 presents an example IDEA assessment, using operating performance data of UK university departments, previously reported in a conventional DEA assessment. Finally, Section 2.5 offers some concluding remarks.

---

[27] Provided full information is available on the production possibilities as well, and in addition that the evaluated input-output vector is measured accurately. See Chapter 3 and 4 for ways of dealing with incomplete production information and measurement error.

[28] As discussed in Section 1.2, many different efficiency measures have been employed. Different models make different assumptions about the organisational objectives. For example, the output-oriented model discussed in Chapter 1 assumes that equiproportionate output augmentation at the current input levels is the objective. Alternatively, input-oriented models assume that equiproportionate input augmentation at the current output levels is the objective. In this respect, models that focus on technical efficiency are 'subjective'. However, if the decision-maker is nonsatiable, any technical efficiency measure gives a necessary condition for effectiveness. Therefore, the models that focus on technical efficiency can be considered 'objective'.

For simplicity, we focus on extending the standard BCC model introduced in the introduction. Therefore, the standard assumptions are maintained, including convexity and accurate measurement. However, we are convinced that the extensions apply with equal strength to alternative DEA models, including those discussed in the subsequent chapters.

2.2. Improving goal congruence

Much research effort has been directed at extending DEA to include additional organisational objectives in addition to technical efficiency. At issue is not so much what assumptions are appropriate, but rather how judgement can be clarified and incorporated in DEA assessments. To incorporate alternative or additional preference information, a number of approaches have been proposed. As displayed in Figure 2.2, we distinguish three alternative approaches for introducing such information:

(1) *behavioural assumptions from micro-economic theory*,

(2) *restrictions formulated by the decision maker*, and

(3) *Interactive decision procedures*.

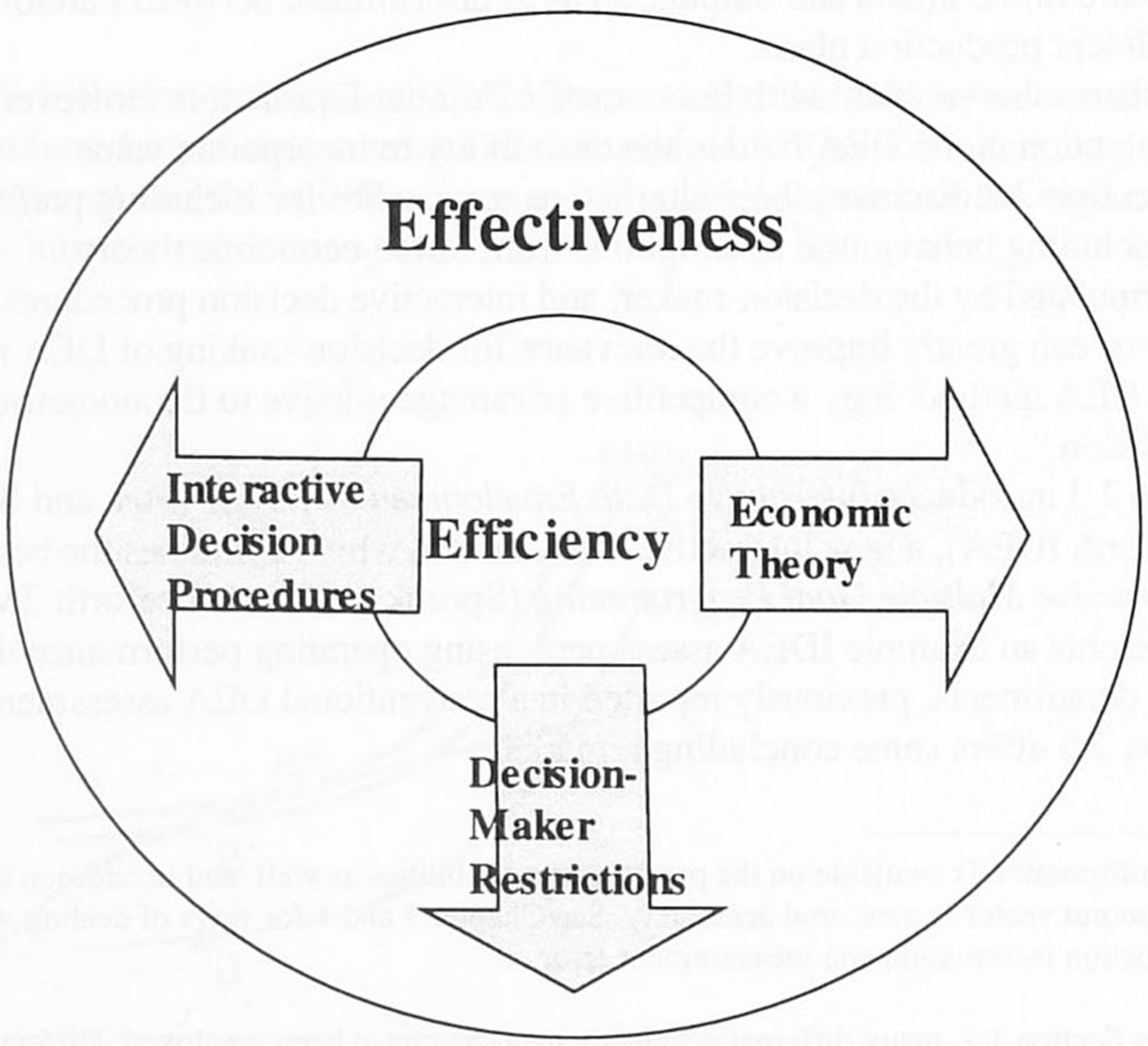


Figure 2. 2 Tools available for including preference information in DEA models

1. Behavioural assumptions from micro-economic theory

A number of DEA studies have derived models that measure performance and select reference units on the basis of behavioural assumptions from micro-economic theory, like cost minimisation, revenue maximisation and profit maximisation.

For example, the neo-classical theory of production postulates that firms maximise profit subject to certain technological constraints. Based on this assumption, Nerlove (1965) introduced the concept of *economic efficiency*, which is defined as the ratio of realised profit to the maximum profit that is achievable given the prices of inputs $w_k = (w_{1k} \cdots w_{mk})^T \in \Re_+^m$ and outputs $p_k = (p_{1k} \cdots p_{pk})^T \in \Re_+^p$ and the technology, i.e.:

$$\textbf{(2. 3)} \quad \rho_k = \frac{p_k^T y_k - w_k^T x_k}{\max_{y,x} \{ p_k^T y - w_k^T x \mid (y,x) \in T \}}.$$

Figure 2.3 illustrates the concept of economic efficiency for the example introduced in Section 1.3. The figure displays two iso-profit lines associated with hypothetical prices for DMU B. The line $p_B^T y - w_B^T x = P_B$ represents the iso-profit line associated with the profit of DMU B, whereas $p_B^T y - w_B^T x = P_{B^*}$. represents the maximum profit achievable given the frontier g. The optimal reference point is B^*, and economic efficiency is defined as $\rho_B = P_B / P_{B^*}$.

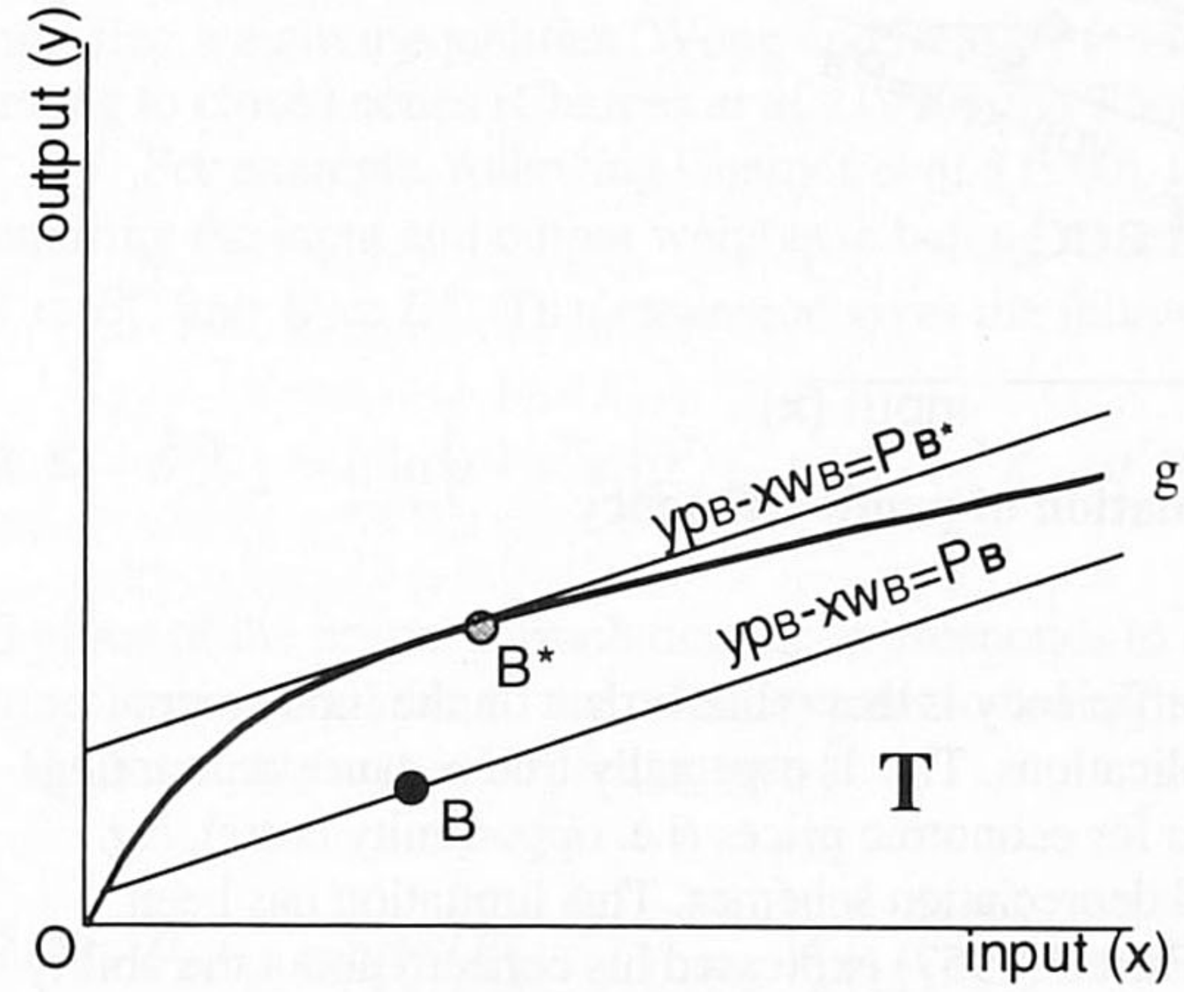


Figure 2. 3 The concept of profit efficiency

In the nonparametric literature, Varian (1984), Banker and Maindiratta (1988), and Färe and Grosskopf (1997), among others, have proposed empirical tests and efficiency measures related to profit efficiency. For example, Banker and Maindiratta (1988) obtained the following upper bound for this measure by replacing the true production set with an inner bound set that corresponds to the BCC set (1.9):

$$\textbf{(2. 4)} \quad \hat{\rho}_k = \frac{p_k^T y_k - w_k^T x_k}{\max_{y,x} \left\{ p_k^T y - w_k^T x \,\middle|\, (y,x) \in \hat{T}_{BCC} \right\}}$$

$$= \frac{p_k^T y_k - w_k^T x_k}{\max_{y,x,\lambda} \left\{ p_k^T y - w_k^T x \,\middle|\, \begin{pmatrix} y \\ -x \end{pmatrix} \leq \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e = 1; \lambda \in \Re_+^n \right\}}.$$

Figure 2.4 illustrates the upper bound for economic efficiency for the example. In addition to the iso-profit line for B, the figure displays $p_B^T y - w_B^T x = P_C$, which represents the iso-profit line associated with the maximum profit achievable given the empirical production set T_{BCC}. The optimal reference point is C, and economic efficiency is estimated as $\hat{\rho}_B = P_B / P_C$.

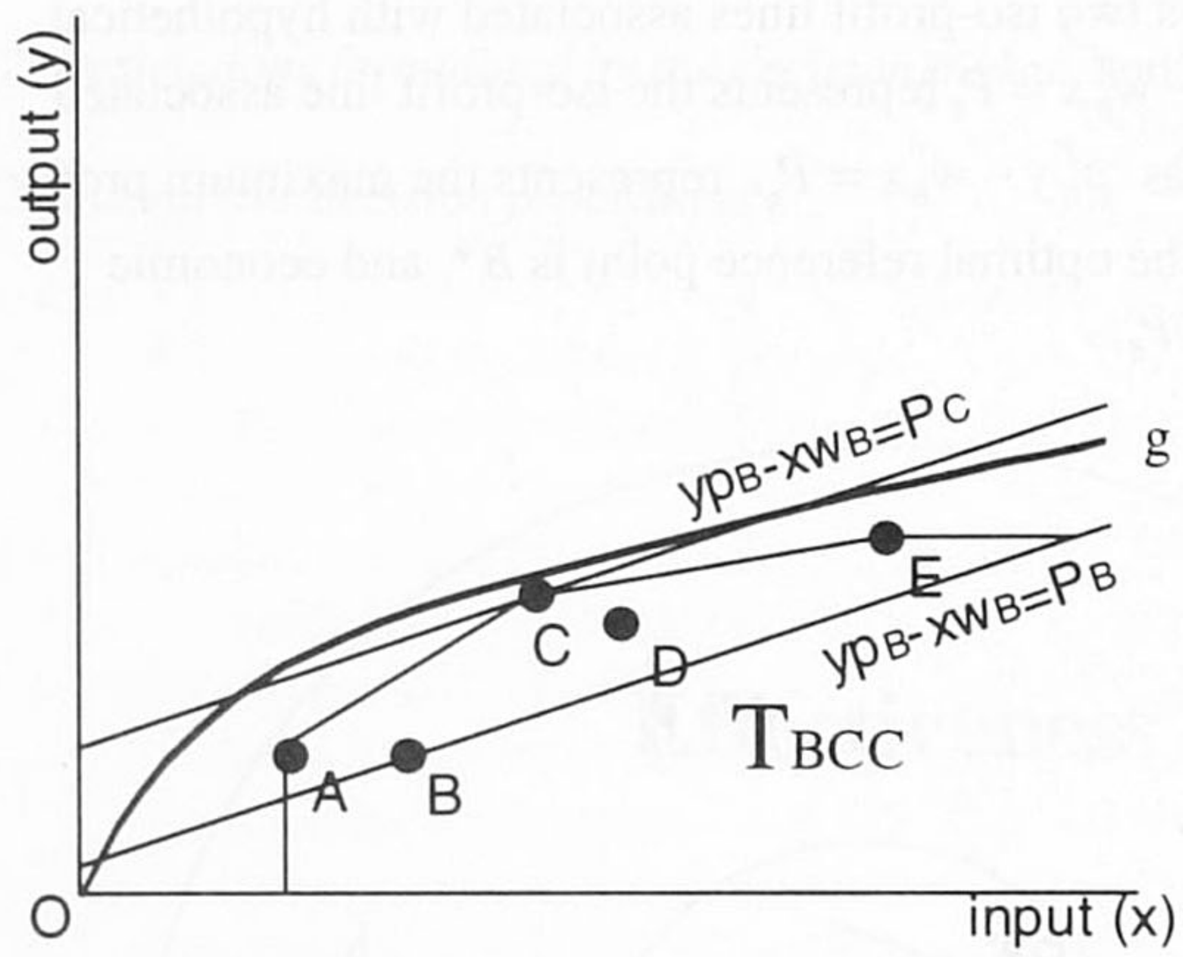


Figure 2. 4 Nonparametric estimation of profit efficiency

A major drawback of using profit efficiency is that reliable data on the input-output prices is not available in many applications. This is especially true because accounting data can give a poor approximation for economic prices (i.e. opportunity costs), e.g. because of debatable valuation and depreciation schemes. This limitation has been recognised in the DEA literature. Farrell (1957) expressed his concern about the ability to measure prices sufficienctly accurate to make good use of allocative efficiency measurement. Charnes and Cooper (1985) cite his concern as a motivation for emphasising technical efficiency. However, a number of solutions have been proposed to deal with incomplete price information. Russell (1985) demonstrated that the output (input) oriented Debreu-Farrell measure provides a lower (upper) bound for revenue (cost) efficiency, i.e. the ratio of maximum revenue (minimum cost) to actual revenue (cost). Building on these results, Kuosmanen and Post (1999c) derived general upper and lower bounds for economic efficiency measures for cases in which price information is limited to particular polyhedral convex cones.

Another problem is that the input-output prices typically involve uncertainty for the DMUs. Prices are typically determined by demand and supply forces that are (partly) beyond the control of the DMUs, and in addition that are (partly) unpredictable for the DMUs. For example, in the classic Cournot (1838) duopoly model, DMUs control their production plans, but prices are determined on the market. Under such circumstances, DMUs require perfect information on demand responses and production plans of competitors, to perfectly anticipate (the Nash equilibrium) prices. Moreover, even if the nominal prices would be fixed i.e. by long-term contracts or regulatory acts, the unpredictable fluctuations in uncontrollable environmental factors such as business cycles, interest rates, and inflation guarantee that the real prices in terms of opportunity costs typically involve uncertainties beyond hedging. Thus, evaluating efficiency of DMUs at the realised prices observed *ex post* by the analyst is unfair towards the units who actually did not have all that information at hand when they fixed their production plans. Kuosmanen and Post (1999b) address this problem. They include uncertainty in DEA by using the theory of stochastic dominance.

2. Restrictions formulated by the decision-maker

Another approach to improving the relevance for decision-making of DEA performance measures and reference units is by including preferences expressed by the decision maker in the objective and/or the constraints of the DEA model.

An approach that has seen considerable application is limiting the flexibility of the model in assigning values to the input-output weights in the dual problem formulation (1.25). Proposed techniques include imposing bounds on individual weights (Dyson and Thanassoulis (1988)) or weight ratios (Thompson *et al.* (1986)), imposing weight inequalities (Wong and Beasley (1990)), and requiring weights to belong to closed cones (Charnes *et al.* (1990) and Thompson *et al.* (1995)).

For example, following Charnes *et al.* (1990), the dual can be extended by requiring the input and output weights to belong to certain closed convex cones, say $V \subset E_+^m$ and $U \subset E_+^p$. That extension gives the following weight restricted model:

$$\textbf{(2. 5)} \quad \hat{\theta}'^{D}_{BCC,k} = \min_{u,v,w} \left\{ u + v^T x_k \,\middle|\, \mu^T y_j = 1; u + v^T X - \mu^T Y \le 0; v^T \in V, \mu^T \in U \right\}$$

In terms of the primal formulation this corresponds to restricting the performance improvements relative to the reference unit to belong to the negative polar cones, i.e. $-V^*$ for inputs and $-U^*$ for outputs[29].

$$\textbf{(2. 6)} \quad \hat{\theta}_{BCC,k} = \max_{\lambda,\theta} \left\{ \theta \,\middle|\, \begin{pmatrix} \theta y_k - Y\lambda \\ -x_k + X\lambda \end{pmatrix} \in \begin{pmatrix} -U^* \\ -V^* \end{pmatrix}; \lambda^T e = 1; \lambda \ge 0 \right\}.$$

For a detailed survey of alternative methods of specifying weight restrictions in DEA, see Allen *et al.* (1997).

Although these methods can considerably increase the relevance for decision-making, they are handicapped by the requirement for substantial amounts of a priori articulated general preference information. Such information is typically difficult to

---

[29] More specifically, if $V = \{A^T \alpha : \alpha \ge 0\}$ and $U = \{B^T \beta : \beta \ge 0\}$, then $-V^* = \{v : Av \le 0\}$ and $-U^* = \{u : B^T u \le 0\}$.

obtain, because it requires considering hypothetical choices and trade-off questions that possibly have little practical appeal. For example, in the single input (output) case, weights have the straightforward interpretation of marginal input (output) levels (see Dyson and Thanassoulis (1988)), but the interpretation of weights in the more general case of multiple inputs and outputs is less intuitively appealing (see Belton (1990)). Moreover, the a priori articulation of managerial preferences increases the information requirement by considering choices and trade-off questions that may prove redundant.

In addition, tight restrictions, such as high aspiration levels for the inputs and outputs or strong weight restrictions may be so tight as to cause infeasible solutions.

3. Interactive decision procedures

To circumvent the problems of using a priori articulated general preference information, it has been suggested to incorporate judgement by combining DEA with interactive decision procedures from Multiple Criteria Decision Making (MCDM) theory. For example, Golany (1988) proposed an interactive decision procedure, combining DEA and Multi-Objective Linear-Programming (MOLP), that allows the decision maker to select a most preferred alternative from a set of alternative DEA efficient points. Belton and Vickers (1992, 1993) incorporated the DEA model within a visual interactive model (Visual Interactive Data Envelopment Analysis, VIDEA), to support decision-makers by displaying the implications of weight restrictions on the efficient frontier and DMU efficiencies. Interactive decision procedures do not require substantial amounts of a priori articulated general preference information, because they rely on the interactive articulation of local preference information, by confronting the decision maker with well-defined and feasible decision alternatives. Local preference information should be easier for the decision-maker to provide and it should reflect more accurately his/her underlying judgement. The choices and trade-offs involved in answering these questions typically have stronger practical appeal than the a priori specification of general preference relationships. Moreover, considering well-defined and feasible decision alternatives only has the additional benefit of reducing the information requirement by avoiding redundant choices and trade-off questions.

An additional benefit of interactive decision procedures is that the decision-maker may become more involved in the process of solving the decision problem, thus obtaining more insight into the decision situation. This embedded learning effect can shape and alter his preferences and may even initiate a restructuring of the model describing the decision problem, by adding omitted input-output variables or by introducing additional DMUs. In addition, the embedded learning effect may help to circumvent imposing restrictions (such as aspiration levels or weight constraints) that are so tight as to cause unfeasible solutions.

To obtain a performance measure, Halme, Joro, Korhonen, Salo and Wallenius (1999) propose the so-called value efficiency approach to incorporating preference information in DEA. That approach first uses an interactive MOLP model to identify the most preferred input-output vector (most preferred solution, or MPS) from the efficient surface. Subsequently, it approximates the indifference contour of the value function at the MPS in terms of its possible tangents, and measures value efficiency relative to that approximation.

2.3. Interactive Data Envelopment Analysis

Although the proposed interactive techniques circumvent the requirement for a priori articulated general preference information, it is unclear whether the interactive decision procedures they embed can reliably elicit preference information from decision makers in complex decision problems (i.e. problems with large numbers of decision alternatives and goal variables). The fact is, empirical results from experimental psychological research indicate that interactive decision procedures that rely on the information processing operations of selecting a most preferred alternative from a set of alternative decision alternatives, or alternatively on assigning criterion weights or weight bounds, can cause decision makers to be inconsistent and to use simplified heuristics (e.g. to drop criteria) in complex decision problem, i.e. problems that involve large numbers of decision alternatives and goal variables (e.g. Larichev *et al.* (1987)). Therefore, we propose to combine DEA and Interactive Multiple Goal Programming (IMGP)[30]. That is an interactive decision procedure that relies on the identification of criterion values that must be improved, may deteriorate, or must remain at least equal to the attained satisfactory value. Psychological research indicates that decision-makers can perform that elementary information processing operation with inconsiderable contradictions and complex strategies (e.g. estimating a set of several criteria values)[31].

An additional argument for proposing IMGP is that the MOLP-DEA and VIDEA approaches consider efficient frontier units only, and, therefore, do not allow the decision maker to select a reference unit from the PPS interior, if, for any reason, he wishes to deliberately include some degree of organisational slack in performance benchmarking. The decision-maker may wish to intentionally include internal slacks, because of organisational and psychological considerations that are excluded from the original model formulation. For example, best observed practice is a performance standard that can not be exceeded (given observed past operating practice and imposed production assumptions[32], and consequently may be stressful and demoralising to DMU managers. By contrast, IMGP does allow the decision-maker to deliberately include some organisational slacks in performance benchmarking, because it considers both production frontier units and PPS interior units. In this Section we combine Data Envelopment Analysis (DEA) and Interactive Multiple Goal Programming (IMGP) into an interactive decision procedure for selecting performance benchmarks, called Interactive Data Envelopment Analysis (IDEA).

The IDEA procedure starts by selecting an initial performance profile (i.e. the collection of input-output levels of some composite DMU) from the production possibility set (PPS), to act as a proposed performance benchmark. Next, the improvements on this performance profile are computed, that are feasible within the DEA PPS. The initial profile and the potential improvements are presented to the

---

[30] For a more detailed discussion of interactive decision procedures and Interactive Multiple Goal Programming, as well as applications in capital budgeting and financial planning, see Nijkamp and Spronk (1980) and Spronk (1981).

[31] In an analytical survey of interactive procedures for multi-criterion linear programming problems, Larichev, Polyakov and Nikiforov (1987) concluded that IMGP is a correct procedure, in terms of reliability of information elicitation from decision makers, sensitivity to random decision maker's error, and speed of convergence to optimal solutions.

[32] Note that DEA efficient reference units need not reflect the theoretical production possibilities, because DEA focuses at observed past operating practice and because it makes only minimal assumptions regarding production relationships. Therefore, DEA efficient performance benchmarks and targets may implicitly include some degree of internal slack, if few observed DMUs operate near the true frontier.

decision maker, who has to indicate whether the proposed performance profile is a satisfactory performance benchmark, and if not, which of the input-output levels should be improved, and to what extent. On the basis of the resulting adjusted performance profile, the constraints on the values of the corresponding input-output variables are reformulated and a new set of potential improvements is calculated and presented to the decision maker, indicating the margins for further adjustments. By successively adjusting the performance profile from iteration to iteration, keeping only the alternatives that meet the higher standards gradually reduces the set of feasible performance profiles, until a satisfactory performance benchmark is identified.

The IDEA performance benchmarking procedure basically involves three main steps:

(1) Selecting a feasible initial performance profile,

(2) (Re-) computing potential performance improvements, and

(3) (Re-) adjusting the selected performance profile.

These steps are subsequently discussed in more detail in the next Sections. Figure 2.5 presents a flow chart that Figureically illustrates the structure of the IDEA performance benchmarking procedure.

To fully exploit all possible techniques for eliciting reliable information from the decision maker, interactive decision procedures can be combined with techniques relying on incorporating prior constraints. To allow for this possibility, we combine the above IMGP procedure with virtual weight constraints. Following Charnes *et al.* (1990), we restrict the dual input-output weights to the closed convex cones V (for inputs) and U (for outputs). This restriction is equivalent to restricting the primal performance improvements to the negative polar cones of $-V^*$ (for inputs) and $-U^*$ (for outputs). The spanning vectors of these polar cones are the directions normal to the hyperplanes bounding V and U.

STEP 1: Selecting a feasible initial performance profile

The first step is the specification of a feasible initial performance profile, by selecting feasible starting values for all input-output variables. This corresponds to selecting some composite DMU from the production possibility set. Obvious candidates for these initial performance levels are the current input-output levels of the DMU under evaluation. Using current performance has the attraction that a feasible solution is assured. Moreover, current performance levels can be readily observed and have strong practical appeal. Alternatively, the decision-maker may short-cut part of the procedure by using a priori specified minimum standards as a starting point for the IDEA procedure. In that case, an initial feasibility check is required, because the aspired performance levels may be outside the production possibility set. Another alternative is using some DEA efficient performance profile, in which case an auxiliary DEA assessment is required to identify the DEA efficient reference unit. Throughout this chapter, the performance standards selected are displayed in capitals, that is, X_{ik} (Y_{ik}) for the i-th (r-th) input (output) variable of the k-th DMU.

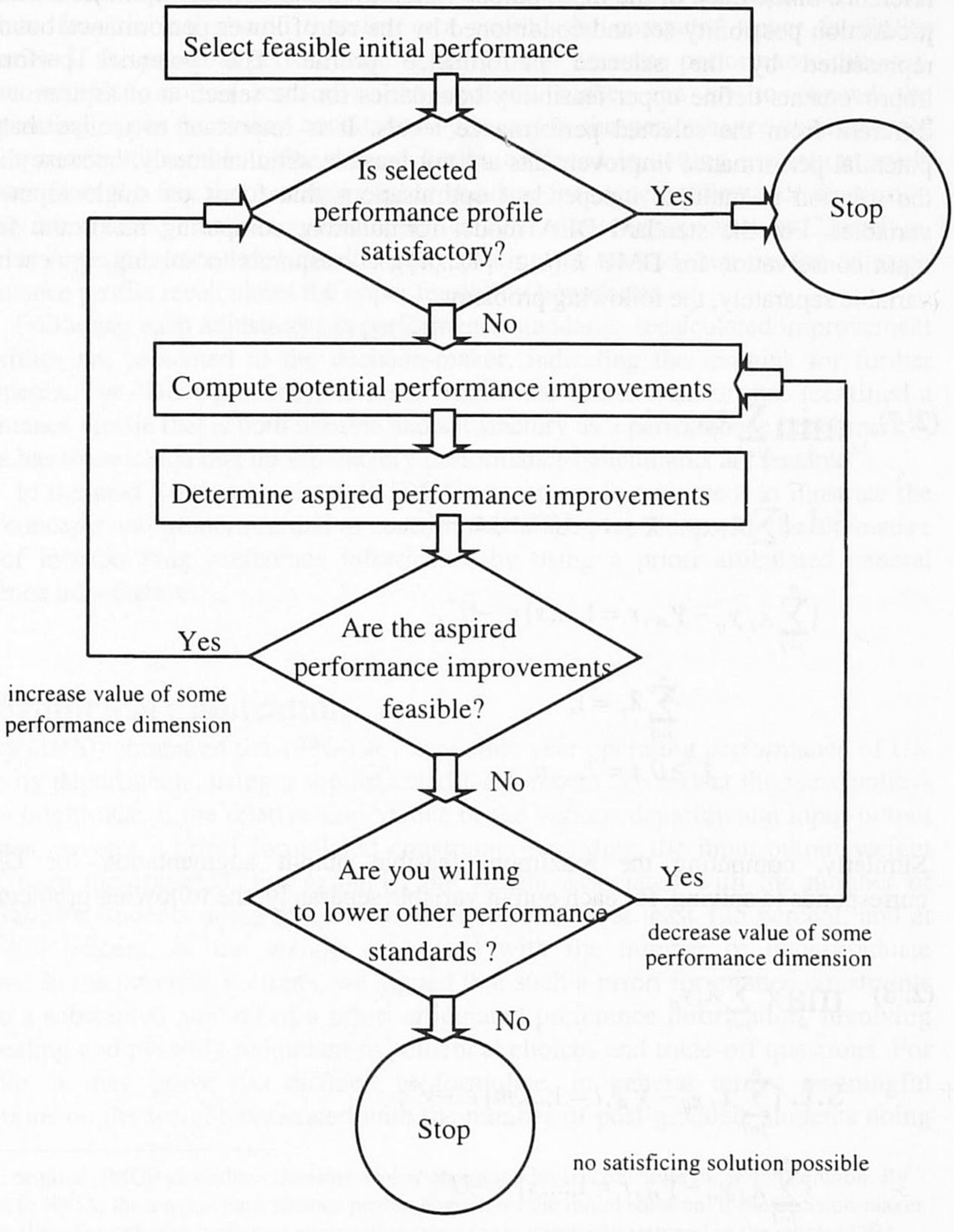


Figure 2. 5 Structure of the IDEA procedure

STEP 2: (Re-) computing potential performance improvements

The second step is the computation, for each input-output variable separately, of the maximum performance improvement that is feasible without deterioration in any other performance dimension. This corresponds to projecting the selected composite DMU in all directions onto the DEA envelopment surface to find multiple DEA efficient reference units. Each of the input-output variables is successively optimised within the production possibility set and conditioned by the set of lower performance boundaries represented by the selected performance profile. The potential performance improvements define upper feasibility boundaries for the selection of aspiration levels different from the selected performance levels. It is important to realise that these potential performance improvements are not feasible simultaneously, because they are the solution to multiple independent optimisations that focus on single input-output variables. For the standard DEA model formulation, computing maximum feasible input conservation for DMU k ($k \in \{1,...,n\}$), corresponds to solving, for each input variable separately, the following problem:

$$(2.7)\quad \min_{\lambda} \sum_{j=1}^{n} \lambda_j x_{ij}$$

$$\text{s.t.} \{\sum_{j=1}^{n} \lambda_j x_{ij} - X_{ik}, i = 1,...,m\} \in -V^*;$$

$$\{\sum_{j=1}^{n} \lambda_j y_{rj} - Y_{rk}, r = 1,...,s\} \in -U^*;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,...,n.$$

Similarly, computing the maximum feasible output augmentation for DMU k corresponds to solving, for each output variable separately, the following problem:

$$(2.8)\quad \max_{\lambda} \sum_{j=1}^{n} \lambda_j y_{rj}$$

$$\text{s.t.} \{\sum_{j=1}^{n} \lambda_j x_{ij} - X_{ik}, i = 1,...,m\} \in -V^*;$$

$$\{\sum_{j=1}^{n} \lambda_j y_{rj} - Y_{rk}, r = 1,...,s\} \in -U^*;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,...,n.$$

STEP 3: (Re-) adjusting the selected performance profile

The third step is the adjustment by the decision-maker of the selected performance profile by sequentially specifying aspired input-output levels. The decision-maker is confronted with the question whether the selected profile is satisfactory, and if not, which performance dimensions should be adjusted, and to what extent. In selecting aspiration levels, the decision maker faces two limitations. First, the aspiration levels should be feasible and should therefore not exceed the computed upper feasibility boundaries. If the potential performance improvements on some performance dimensions are regarded as too restrictive, the decision maker can try to expand the corresponding upper feasibility boundaries by lowering the performance standards for other input-output variables. The second limitation faced in setting aspiration levels originates from the fact that setting a higher standard for some performance dimension typically lowers the potential improvements for other input-output variables. Therefore, whenever a performance level is adjusted, repeating the second step for the adjusted performance profile recalculates the upper feasibility boundaries.

Following each adjustment in performance standards, recalculated improvement possibilities are presented to the decision-maker, indicating the margins for further adjustments. The IDEA procedure is iterated until the decision-maker has identified a performance profile that is both feasible and satisfactory as a performance benchmark[33], or if he has to conclude that no satisfactory performance benchmarks are feasible[34].

In the next Section an example IDEA assessment is presented, to illustrate the IDEA concepts and properties, and to contrast the IDEA procedure with the alternative route of incorporating preference information by using a priori articulated general preference information.

2.4. Empirical application

Beasley (1990) compared the 1986-1987 academic year operating performance of UK university departments, using a sophisticated DEA model. To reflect the view policy-makers might take at the relative importance of the various departmental input-output measures, several a priori formulated constraints regarding the input-output weight values were incorporated. For example, the weight associated with the number of post-graduate students doing research was required to be at least 125 percent, and at most 200 percent, of the weight associated with the number of undergraduate students. In the previous sections, we argued that such a priori formulated constraints require a substantial amount of a priori articulated preference information, involving unappealing and possibly redundant hypothetical choices and trade-off questions. For example, it may prove too difficult to formulate, in general terms, meaningful restrictions on the weight associated with the number of post-graduate students doing

---

[33] In the original IMGP procedure, decision-maker aspiration levels constitute the initial solution. By contrast, in IDEA, the current performance profile constitutes the initial solution. If the decision-maker does not allow for reducing output or augmenting input (as is implicitly assumed in the current DEA models), this effectively increases the rate of convergence. By contrast, if output reductions and input augmentations are allowed, starting from the current performance profile can reduce the rate of convergence.

[34] Failure to find a satisfactory benchmark can mean that the objectives are simply not attainable. An alternative explanation is that the empirical production set gives an incomplete representation of the production possibilities, due to sampling error. This problem can be remedied by including additional observations or additional production assumptions (see Chapter 3).

research, because its relative importance may depend on the current student population composition[35]. In addition, a priori formulated constraints may prove so tight as to cause unfeasible solutions, and do not allow for the deliberate inclusion of organisational slacks. To contrast the IDEA procedure with the alternative approach of incorporating managerial preferences by using a priori formulated constraints, this section presents an example IDEA assessment using the physics departments data set from the Beasley study.

Description of the IDEA model

The data set contains seven input-output variables referring to the 1986-1987 academic year operating performance of 50 physics departments of UK universities. The complete data set employed is displayed in Appendix 2.1[36]. The departmental input-output variables used in the analysis are[37]:

Inputs: x_1 the amount of general expenditure (x £1000)

x_2 the amount of equipment expenditure (x £1000)

Outputs: y_1 the amount of research income (x £1000)

y_2 the number of undergraduate students

y_3 the number of post-graduate students on taught courses

y_4 the number of post-graduate students doing research

y_5 the University Grant Committee research rating

The UGC research rating is a categorical variable, that can take on any of four values: (1) a below average rating, (2) an average rating, (3) an above average rating or (4) a star rating. It can be computed as the sum of four dummy variables that are assigned a unity value if the research rating of a department is the same or more favourable than the rating the dummy variables represent, and else are assigned a zero value. The research rating dummy variables employed are:

w_1 below-average rating dummy

w_2 average rating dummy

---

[35] Stewart (1996) demonstrates that one of the serious problems in constraining weight flexibility in DEA models is the potential violation of additive difference independence, i.e. the property that increments in inputs (outputs) must represent the same marginal contribution to the virtual input (output), no matter what their base lines, or the levels of any other inputs (outputs).

[36] Source: Imperial College Operational Research Library (ftp://graph/ms.ic.ac.uk/ pub/).

[37] The original article stressed that the variable set used may not completely capture the departmental operating performance, because information on the number of publications and citations and indicators of teaching quality were unavailable, and because current expenses will affect research output in future years. For a more detailed discussion of the original data set composition and model specification, see Beasley (1990).

w_3 above-average rating dummy

w_4 star rating dummy

Apart from applying an interactive decision procedure instead of applying a priori formulated restrictions, we deviate from the original DEA assessment in three respects. First, we use the research income variable as an output variable, so as to reflect the success of a department in attracting research projects. In the original model formulation, the research income variable was also used as an input variable, so as to reflect the esteem in which a department, and its work, is held.

Second, we incorporate the UGC research rating variable as a categorical variable, instead of incorporating it as a continuous variable. To do so, we partitioned the department sample in four research rating category sub-samples:

D_1 below-average rating subset

D_2 average rating subset

D_3 above-average rating subset

D_4 star rating subset

These categories are exhaustive and mutually exclusive and have a natural hierarchy, that is, star is more favourable than above average, above-average is more favourable than average, and average is more favourable below-average. In the computation of the potential performance improvements for the continuous input-output variables (x_1, x_2, y_1, y_2, y_3 and y_4) we add to the basic model the constraint that all departments included in the composite reference unit are necessarily of the same or more favourable research rating category as the department being analysed (see Charnes, Cooper, Lewin, and Seiford (1994)). For example, a department rated above average can be compared with other departments rated above average and with departments rated star. In the computation of the potential performance improvements for the categorical research rating variable we can not use this method, because it may generate non-integer solutions. Therefore, we use a mixed-integer LP model formulation to compute feasible research rating improvements (see Banker and Morey (1986)).

Third, the original model used the Charnes, Cooper and Rhodes (1978) model that assumed ray unboundedness in addition to convexity and free disposability (see Chapter 3). By contrast, we use the BCC model discussed in Chapter 1 that assume convexity and free disposability only.

The resulting model for computing the maximum feasible input conservation for department k ($k \in \{1,...,n\}$), corresponds to problem :

$$(2.9)\quad \min_{\lambda,s,s'} \sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j x_{ij}$$

$$\text{s.t.} \sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j y_{rj} - s'_r = Y_{rk} \quad r=1,2,3,4;$$

$$\sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j x_{ij} + s_r = X_{ik} \quad i=1,2;$$

$$\sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j\in \bigcup_{k=y_{sk}}^{4} D_t;$$

$$s'_r \geq 0 \quad r=1,2,3,4;$$

$$s_i \geq 0 \quad i=1,2.$$

By analogy, the resulting model for computing the maximum feasible output augmentation corresponds to problem:

$$(2.10)\quad \max_{\lambda,s,s'} \sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j y_{rj}$$

$$\text{s.t.} \sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j y_{rj} - s'_r = Y_{rk} \quad r=1,2,3,4;$$

$$\sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j x_{ij} + s_r = X_{ik} \quad i=1,2;$$

$$\sum_{j\in \bigcup_{t=y_{sk}}^{4} D_t} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j\in \bigcup_{k=y_{sk}}^{4} D_t;$$

$$s'_r \geq 0 \quad r=1,2,3,4;$$

$$s_i \geq 0 \quad i=1,2.$$

Similarly, the resulting model for computing the maximum feasible research rating improvement corresponds to the following, using W_{lj} to denote the aspired level for l-th research rating dummy of j-th department:

$$(2.11)\quad \max_{\lambda,s,s's,s'} \sum_{l=1}^{4} s_l''$$

$$\text{s.t.} \sum_{j=1}^{50} \lambda_j y_{rj} - s_r' = Y_{rk} \quad r = 1,2,3,4;$$

$$\sum_{j=1}^{50} \lambda_j x_{ij} + s_i = X_{ik} \quad i = 1,2;$$

$$\sum_{j=1}^{50} \lambda_j w_{lj} - s_l'' = W_{lk} \quad l = 1,2,3,4;$$

$$s_l'' \in \{0,1\} \quad l = 1,2,3,4;$$

$$s_{l-1}'' \geq s_l \quad l = 2,3,4;$$

$$\sum_{j=1}^{50} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j=1,\ldots,50;$$

$$s_l'' \geq 0 \quad l=1,2,3,4;$$

$$s_r' \geq 0 \quad l=1,2,3,4;$$

$$s_i \geq 0 \quad i=1,2.$$

An example IDEA assessment

One of the results of the original weight constrained study was that the Birmingham department was highly inefficient, and put relatively much emphasis on the number of undergraduate students. This was reflected in a 58 percent efficiency score and a 60 percent virtual output level for the number of undergraduates[38]. The efficient reference unit selected was a multiple of the Essex department. As stated earlier, the problem with using this reference unit as a performance benchmark, is that it may not properly or fully reflect the policy-maker's preferences, because preference information was incorporated by using a priori formulated weight restrictions. To illustrate the IDEA procedure, and, moreover, to contrast it with the weight constrained approach, we conducted an IDEA assessment for the Birmingham department. The output information of this assessment is displayed in Table 2.1. We started by selecting Birmingham's current performance as the starting point. As stated earlier, using current performance assures a feasible solution, and, moreover, current performance levels can be readily observed and have strong practical appeal. Next, we computed, for each input-output variable separately, the maximum improvement that is feasible without deterioration in any of the other input-output variables. This first-run operation amounts to solving the mathematical programming problems 5 (for inputs), 6 (for outputs) and 7 (for the research rating), with $X_{ik} = x_{ik}$ (i=1,2), $Y_{rk} = y_{rk}$ (r=1,2,3,4), $W_{lk} = w_{lk}$ (l=1,2,3,4).

[38] A virtual output is the product of an output level and the corresponding weight value. It represents the proportion the output variable represents in the aggregate weighed output, and reflects the dependency of the DEA efficiency score on the output variable.

Performance Dimension	Run	Selected Performance Profile	Upper Feasibility Boundary	Potential Performance Improvements
General Expenditure	1	2605	1981	624
	2	2700	1981	719
	3	2700	2619	81
Equipment Expenditure	1	301	171	130
	2	301	171	130
	3	301	200	101
Research Income	1	1485	3185	1700
	2	1485	3342	1857
	3	1485	2513	1028
Undergraduate Students	1	381	504	123
	2	381	552	171
	3	381	392	11
Post-graduates on Taught Courses	1	16	79	63
	2	16	79	63
	3	16	17	1
Post-graduates doing Research	1	54	118	64
	2	54	134	69
	3	54	123	69
UGC Research Rating	1	3	3	0
	2	4	4	1
	3	4	4	0

Table 2. 1 IDEA results for the Birmingham department for 3 runs

The computed improvement possibilities can be used to select performance benchmarks, depending on organisational policy considerations, managerial preferences and external restrictions. In the weight-constrained study, the author suggested that Birmingham should achieve a star research rating, and should double its number of post-graduate students doing research. In our analysis we adhered to these suggestions. The output table shows that the aspired doubling of the number of postgraduates doing research is feasible without any concession on other performance dimensions. By contrast, the aspired star research rating cannot be achieved without lowering the performance standards for other input-output variables. Therefore, we assumed that the policy-makers are prepared to increase general expenses from the current £2,605,000 level to £2,700,000 (possibly by appointing a distinguished professor). Using the adjusted input-output levels, potential improvements were recalculated (run 2). These computations indicate that by increasing the amount of general expenditure by £95,000, the aspired star research rating becomes feasible. The above findings imply that a constrained DEA model incorporating the a priori formulated restriction that a star

research rating is achieved, would have had no feasible solution for the Birmingham department. By contrast, the IDEA procedure circumvents unfeasible solutions by computing, at each iteration, upper feasibility boundaries before aspiration levels are specified. Next, we assumed that a star research rating is indeed selected as an aspiration level. This induces the run 3 improvement possibilities. These potential improvements indicate that the aspired doubling of the number of post-graduate students doing research is feasible without making any further concessions. In fact, any number up to 123 postgraduates could be selected as an aspiration level. The policy-makers could continue this procedure, so as to fine-tune the Birmingham performance standards, depending on their insights and preferences, or to explore the PPS to get additional insights into its characteristics. However, we assumed that the policy-makers are being satisfied with the performance profile that results if Birmingham achieves a star rating and doubles its number of post-graduates doing research, while keeping all other input-output variables at their current levels. This implies that a considerable amount of internal slack is incorporated in the resulting performance standards, because the performance possibilities are not completely exploited. The remaining performance improvement possibilities are listed in the table below. That table also lists the departments included in the composite reference units that correspond to the optimisations of the various performance dimensions. The London Imperial and Surrey departments are referencing units in all of these optimisations. These departments are potential benchmark partners that may be helpful in uncovering routes for achieving the selected performance standards.

Performance Dimension	Selected Performance Profile	Potential Performance Improvements	Benchmark Partners
General Expenditure	2700	81	London Imperial, Surrey
Equipment Expenditure	301	101	London Imperial, Surrey
Research Income	1485	1028	London Imperial, Surrey, Cambridge
Undergraduate Students	381	11	London Imperial, Surrey
Post-Graduates Taught Courses	16	1	London Imperial, Surrey, Manchester
Post-Graduates doing Research	108	15	London Imperial, Surrey, Cambridge
UGC Research Rating	4	0	London Imperial, Surrey

Table 2. 2 Final results of the IDEA procedure for the Birmingham department

2.5. Concluding remarks

The original DEA models focus on the evaluation of technical efficiency. However, since technical efficiency generally is not the sole organisation objective, there is a need to include additional information on the desirable structure of the inputs and the outputs. This chapter demonstrated that the DEA methodology contains a number of powerful tools for including such information, including behavioural assumptions from micro-economic theory, restrictions formulated by the decision-maker, and interactive decision procedures. These tools can greatly improve the relevance for decision-making of DEA output. This gives the DEA methodology a competitive advantage relative to Econometric Frontier Estimation, which does not involve much flexibility to include preference information in addition to efficiency.

In addition, this chapter introduced Interactive Data Envelopment Analysis (IDEA), a new interactive performance benchmarking procedure IDEA that combines DEA and the well-established interactive decision procedure Interactive Multiple Goal Programming (IMGP). Interactive decision procedures are powerful tools, because they do not require the articulation of general preference information prior to the analysis. In addition, interactive decision procedures, can increase the involvement of the decision-maker in solving the decision problem, and increase his/her insight into the decision situation. Furthermore, the embedded learning effect may help to circumvent imposing restrictions that are so tight as to cause unfeasible solutions.

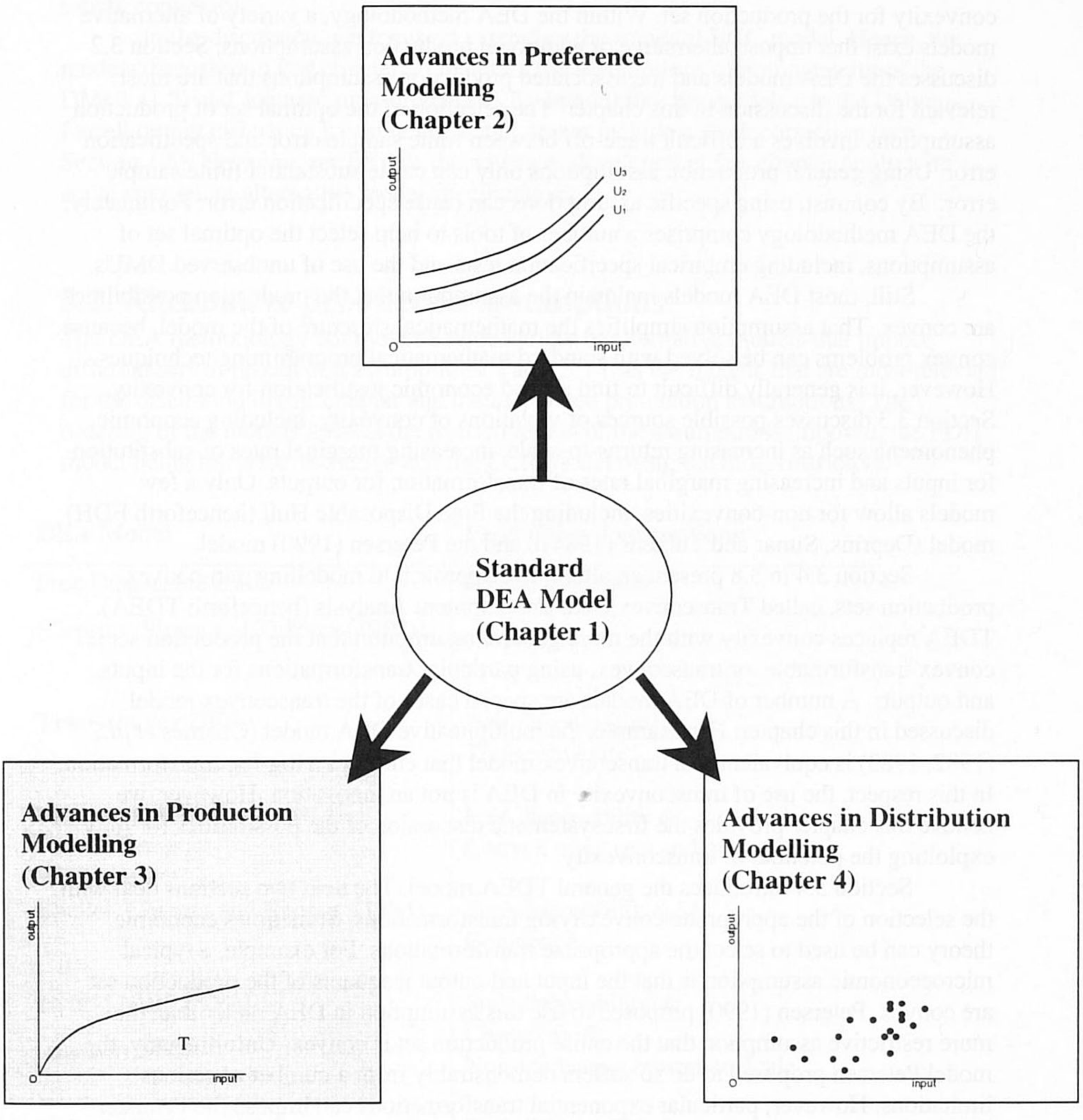
IDEA has a number of advantages relative to the interactive decision procedures proposed in DEA thusfar. These interactive models invariably rely on the information processing operations of selecting a most preferred alternative from a set of alternative decision alternatives, or alternatively on assigning criterion weights or weight bounds. Empirical results from experimental psychological research indicate that interactive decision procedures that rely on these information processing can cause decision makers to be inconsistent and to use simplified heuristics (e.g. to drop criteria) in complex decision problems, i.e. problems that involve large numbers of decision alternatives and goal variables. By contrast, IDEA relies on the identification of criterion values that must be improved, may deteriorate, or must remain at least equal to the attained satisfactory value. Psychological research indicates that decision-makers can perform that elementary information processing operation with inconsiderable contradictions and complex strategies (e.g. estimating a set of several criteria values).

In addition, the interactive models currently available invariably consider production frontier units only, and, therefore, do not allow the decision maker to select a reference unit from the PPS interior if, for any reason, he wishes to deliberately include some degree of organisational slack in performance benchmarking. The decision-maker may wish to intentionally include internal slacks, because of organisational and psychological considerations that are excluded from the original model formulation. For example, best observed practice is a performance standard that can not be exceeded (given observed past operating practice and imposed production assumptions), and, consequently, may be stressful and demoralising to DMU managers. By contrast, IDEA does allow the decision-maker to deliberately include some organisational slacks in performance benchmarking, because it considers both production frontier units and PPS interior units.

Appendix 2.1 The data set for the empirical application.

UNIVERSITY	GEN. EXP. (£1000)	EQUIP. EXP. (£1000)	RES. INC. (£1000)	UGs	PGsT	PGsR	UGC RES. RAT.
BATH	528	64	254	145	0	26	1
BIRMINGHAM	2605	301	1485	381	16	54	3
BRADFORD	304	23	45	44	3	3	1
BRISTOL	1620	485	940	287	0	48	3
BRUNEL	490	90	106	91	8	22	1
CAMBRIDGE	2675	767	2967	352	4	166	4
CITY	422	0	298	70	12	19	1
DURHAM	986	126	776	203	0	32	2
EAST ANGLIA	523	32	39	60	0	17	4
ESSEX	585	87	353	80	17	27	3
EXETER	931	161	293	191	0	20	1
HULL	1060	91	781	139	0	37	1
KEELE	500	109	215	104	0	19	1
KENT	714	77	269	132	0	24	1
LANCASTER	923	121	392	135	10	31	2
LEEDS	1267	128	546	169	0	31	2
LEICESTER	891	116	925	125	0	24	3
LIVERPOOL	1395	571	764	176	14	27	3
LONDON BIRBECK	990	83	615	28	36	57	1
LONDON IMPERIAL	3512	267	3182	511	23	153	4
LONDON KQC	1451	226	791	198	0	53	2
LONDON Q MARY	1018	81	741	161	5	29	3
LONDON R HOL BED	1115	450	347	148	4	32	1
LONDON UNIV COLL	2055	112	2945	207	1	47	3
LOUGHBOROUGH	440	74	453	115	0	9	2
MANCHESTER	3897	841	2331	353	28	65	4
UMIST	836	81	695	129	0	37	1
NEWCASTLE	1007	50	98	174	7	23	2
NOTTINGHAM	1188	170	879	253	0	38	2
OXFORD	4630	628	4838	544	0	217	4
READING	977	77	490	94	26	26	2
SALFORD	829	61	291	128	17	25	1
SHEFFIELD	898	39	327	190	1	18	1
SOUTHAMPTON	901	131	956	168	9	50	3
SURREY	924	119	512	119	37	48	3
SUSSEX	1251	62	563	193	13	43	2
WARWICK	1011	235	714	217	0	36	3
YORK	732	94	297	151	3	23	2
ABERYSTWYTH	444	46	277	49	2	19	2
BANGOR	308	28	154	57	0	7	1
CARDIFF	483	40	531	117	0	23	1
SWANSEA	515	68	305	79	7	23	1
ABERDEEN	593	82	85	101	1	9	1
DUNDEE	570	26	130	71	20	11	1
EDINBURGH	1317	123	1043	293	1	39	2
GLASGOW	2013	149	1523	403	2	51	3
HERIOT-WATT	992	89	743	161	1	30	2
ST ANDREWS	1038	82	513	151	13	47	2
STIRLING	206	1	72	16	0	6	1
STRATHCLYDE	1193	95	485	240	0	32	1

Chapter 3:
Advances in Production Modelling



3.1. Introduction

As discussed before, DEA relies on general production assumptions rather than a specific functional form for the production relationships. This feature makes it possible for the data to 'speak for themselves' rather than being forced to use the idiom of some imposed functional form. This gives DEA a comparative advantage relative to econometric models that typically do require a parametric specification.

The standard BCC model discussed in Chapter 1 assumes free disposability and convexity for the production set. Within the DEA methodology, a variety of alternative models exist that impose alternative or additional production assumptions. Section 3.2 discusses the DEA models and the associated production assumptions that are most relevant for the discussion in this chapter. The selection of the optimal set of production assumptions involves a difficult trade-off between finite sample error and specification error. Using general production assumptions only can cause substantial finite sample error. By contrast, using specific assumptions can cause specification error. Fortunately, the DEA methodology comprises a number of tools to help select the optimal set of assumptions, including empirical specification tests and the use of unobserved DMUs.

Still, most DEA models maintain the assumption that the production possibilities are convex. That assumption simplifies the mathematical structure of the model, because convex problems can be solved with standard mathematical programming techniques. However, it is generally difficult to find a good economic justification for convexity. Section 3.3 discusses possible sources of violations of convexity, including economic phenomena such as increasing returns-to-scale, increasing marginal rates of substitution for inputs and increasing marginal rates of transformation for outputs. Only a few models allow for non-convexities, including the Free Disposable Hull (henceforth FDH) model (Deprins, Simar and Tulkens (1984)), and the Petersen (1990) model.

Section 3.4 to 3.8 present an alternative approach to modelling non-convex production sets, called Transconvex Data Envelopment Analysis (henceforth TDEA). TDEA replaces convexity with the more general assumption that the production set is convex transformable, or transconvex, using particular transformations for the inputs and outputs. A number of DEA models are special cases of the transconvex model discussed in this chapter. For example, the multiplicative DEA model (Charnes *et al.*, (1982, 1983) is equivalent to a transconvex model that employs a log-log transformation. In this respect, the use of transconvexity in DEA is not an innovation. However, we believe this chapter provides the first systematic discussion of the possibilities for fully exploiting the potential of transconvexity.

Section 3.4 introduces the general TDEA model. The next two sections deal with the selection of the appropriate convexifying transformations. Sometimes economic theory can be used to select the appropriate transformations. For example, a typical microeconomic assumption is that the input and output isoquants of the production set are convex. Petersen (1990) proposed to use this assumption in DEA rather than the more restrictive assumption that the entire production set is convex. Unfortunately, the model Petersen proposed to do so suffers demonstrably from a number of serious limitations. However, particular exponential transformations can impose the Petersen assumption in the TDEA approach. Section 3.5 discusses this special case of the TDEA. Apart from economic theory, empirical selection procedures can provide guidance for selecting the appropriate transformations. Section 3.6 develops a specification test based on exponential transformations and a F-test proposed by Banker (1993). Interestingly, that procedure induces a gradual transition between the BCC model and the less restrictive FDH model, models that represent the limiting cases for the exponential transformation.

If the transformations convexify the production set, the TDEA model can be demonstrated to be statistically consistent. However, statistical consistency relates to large sample performance only. Therefore, Section 3.7 simulates the finite sample performance of TDEA using Monte-Carlo simulations. Section 3.8 presents an empirical application for the bank branch network of a large Dutch financial institution. Theoretical and practical arguments, as well as empirical evidence suggest that the bank branch technology is characterised by increasing returns-to-scale and hence does not satisfy convexity.

In the discussion, we focus on extending the standard BCC model. Hence, the models discussed in this chapter all maintain the assumptions of homogeneity of the DMUs (1.5) and accurate measurement (1.8). In addition, the models use the Debreu-Farrell output efficiency measure (1.3), and do not include a slack correction (see Section 1.5). However, we believe the advances developed in this chapter apply with equal strength to alternative model specifications.

3.2. Alternative production assumptions

The DEA methodology comprises a wide variety of alternative models that impose different sets of production assumptions. Table 3.1 lists the models that are most relevant for the discussion in this chapter and the associated production assumptions. The ordering of the models reflects the restrictiveness of the assumptions imposed, the FDH model being the least restrictive and the CCR model being the most restrictive.

DEA Model	**Production Assumptions**
Free Disposable Hull (Deprins, Simar and Tulkens (1984))	Free disposability
Transconvex DEA	Free Disposability & Transconvexity
Petersen (1990)	Free disposability & Convex input sets and output sets
Banker, Charnes and Cooper (1984)	Free disposablity & Convexity
Hybrid NIRS model (Banker 1984)	Free disposability & Convexity & Non-increasing returns-to-scale
Hybrid NDRS model (Banker 1984)	Free disposability & Convexity & Non-decreasing returns-to-scale
Charnes, Cooper and Rhodes (1978)	Free disposability & Convexity & Constant returns-to-scale

Table 3.1 Some important DEA models and the associated production assumptions.

Some DEA models make minimal production assumptions only. For example, the Free Disposable Hull model (Deprins, Simar and Tulkens (1984), see also Tulkens (1993)) assumes free disposability (1.6) only. The associated empirical production set is the following set:

$$\textbf{(3.1)} \quad \hat{T}_{FDH} = \left\{ (y,x) \in \Re_+^{p+m} \middle| \begin{pmatrix} y \\ -x \end{pmatrix} \leq \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e = 1; \lambda_j \in \{0,1\} \quad j \in D \right\}.$$

This set essentially represents all observations in the data set and all input-output vectors that are Pareto-Koopmans dominated by the observations. Measuring efficiency relative to this set gives the following estimator for efficiency:

$$\textbf{(3. 2)} \quad \hat{\theta}_{FDH,k} = \max_{\theta} \left\{ \theta \middle| (\theta y_k, x_k) \in \hat{T}_{FDH} \right\}$$

$$= \max_{\lambda,\theta} \left\{ \theta \middle| \begin{pmatrix} \theta y_k \\ x_k \end{pmatrix} \leq \begin{pmatrix} Y\lambda \\ -X\lambda \end{pmatrix}; \lambda^T e = 1; \lambda_j \in \{0,1\} \quad j \in D \right\}.$$

This model can be solved using standard Mixed Integer Linear Programming techniques[39].

The FDH model restricts the BCC model (1.10) by adding the restriction that the individual lambda values be either zero or unity. Consequently, in this model the evaluated unit can be compared with non-composite reference units only.

Figure 3.1 displays the FDH results for the example data set introduced in Section 1.3. The staircase line f_{FDH} is the FDH frontier. That line connects all input-output vectors that are not Pareto-Koopmans dominated by observations in the data set.

---

[39] Alternatively, the problem can also be solved by enumeration. This is because there necessarily is an observed DMU that gives the optimal solution. Therefore, the solution space can harmlessly be reduced to the set of observed DMUs. Specifically, the following formulation gives the same solution as (3.2):

$$\hat{\theta}_{FDH,k} = \max_j \left\{ \min_r \left\{ \frac{y_{rj}}{y_{rk}} \right\} \middle| x_j \geq x_k \right\}.$$

We used this approach for computing the FDH estimates in the simulations in Section 3.7. Appendix 3.1 presents the associated computer program in GAUSS code.

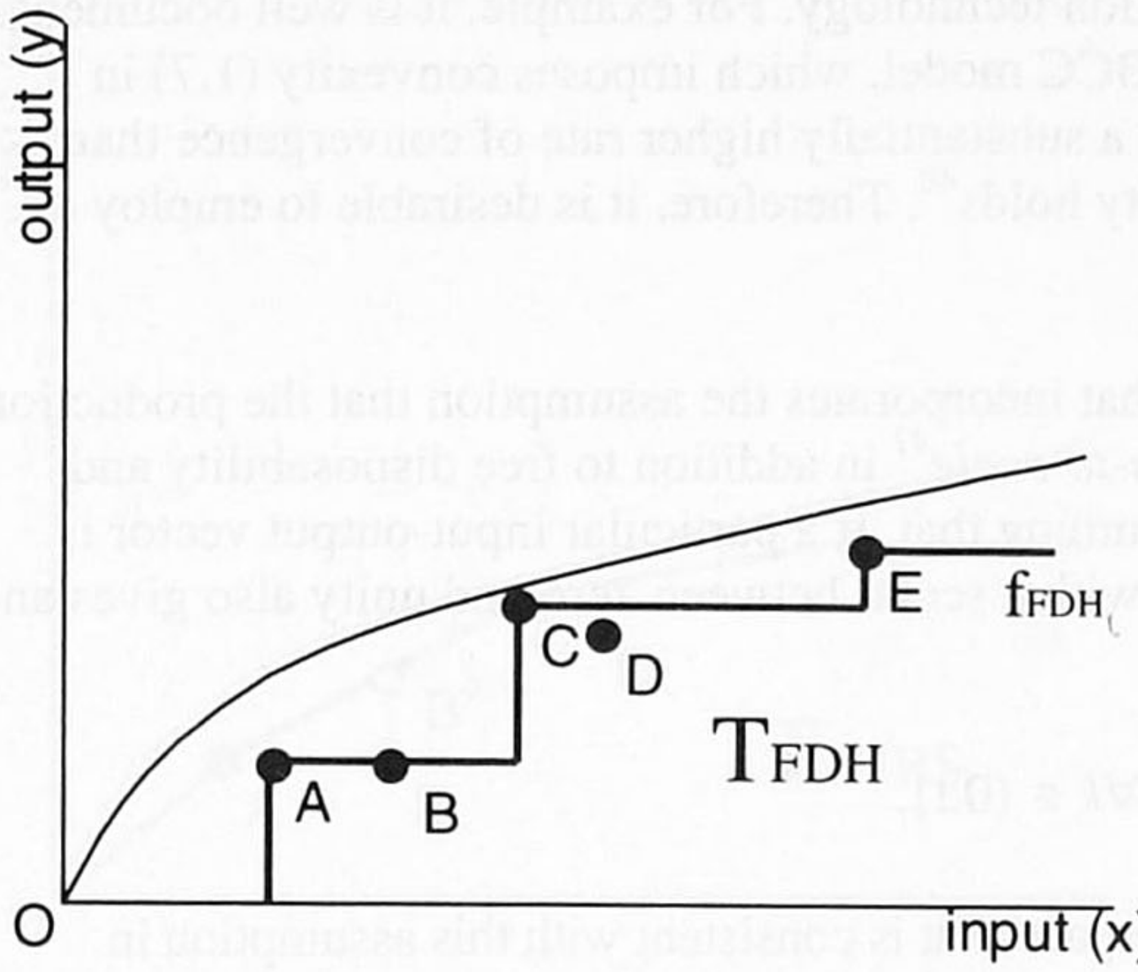


Figure 3. 1 The FDH model for the example data set introduced in Section 1.3

This model gives statistically consistent estimates for all technologies that satisfy free disposability, under the standard distribution assumptions (1.5), (1.8) and (1.15).

An additional advantage of FDH is that the FDH reference units are observed DMUs. By contrast, BCC reference units can be fictitious DMUs, constructed as convex combinations of observations. For DMU managers, fictitious DMUs can be less acceptable as a benchmark than observed DMUs. This consideration is especially important if DEA is used for the purpose of setting targets and selecting reference units in managerial applications. Still, this consideration is less important for the purpose of efficiency estimation.

Unfortunately, models with little prior structure can suffer from substantial *finite sample error* and converge relatively slow, especially if the number of observations is small relative to the number of input and output variables. Since the FDH set is contained within the BCC set, i.e. $\hat{T}_{FDH} \subseteq \hat{T}_{BCC}$, we find that the FDH estimator necessarily falls below the BCC estimator, i.e.:

(3. 3) $\hat{\theta}_{FDH,k} \leq \hat{\theta}_{BCC,k}$.

If the BCC assumptions are satisfied, both estimators necessarily fall below true efficiency, and hence FDH involves more error than BCC:

(3. 4) $\left|\hat{\theta}_{FDH,k} - \theta_k\right| \geq \left|\hat{\theta}_{BCC,k} - \theta_k\right|$.

In Figure 3.1, this is reflected by the fact that FDH wrongly classifies DMU *B* as efficient. Although input-output levels that dominate DMU *B* are attainable, the data set does not contain DMUs that operate at such input-output levels, and hence efficiency is overestimated.

Introducing additional production information can substantially improve finite sample performance. More specific DEA models impose additional assumptions with

respect to the nature of the production technology. For example, it is well documented that efficiency estimates from the BCC model, which imposes convexity (1.7) in addition to free disposability, have a substantially higher rate of convergence than FDH estimators, provided convexity holds[40]. Therefore, it is desirable to employ a larger subset of the production set.

Banker (1984) proposed a model that incorporates the assumption that the production set satisfies *non-increasing returns-to-scale*[41] in addition to free disposability and convexity. This corresponds to assuming that, if a particular input-output vector is attainable, multiplying that vector with a scalar between zero and unity also gives an attainable vector, i.e.:

$$\textbf{(3. 5)} \quad (y,x)\in T \Rightarrow (ky,kx)\in T \quad \forall k\in(0,1].$$

The smallest subset in input-output space that is consistent with this assumption in addition to the BCC assumptions is the following empirical production set:

$$\textbf{(3. 6)} \quad \hat{T}_{NIRS}=\left\{(y,x)\in\Re_{+}^{p+m}\left|\begin{pmatrix}y\\-x\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix};\lambda^{T}e\le 1;\lambda\in\Re_{+}^{n}\right.\right\}.$$

Measuring efficiency relative to this set gives the following estimator:

$$\textbf{(3. 7)} \quad \hat{\theta}_{NIRS,k}=\max_{\theta}\left\{\theta\left|(\theta y_k,x_k)\in\hat{T}_{NIRS}\right.\right\}$$

$$=\max_{\lambda,\theta}\left\{\theta\left|\begin{pmatrix}\theta y_k\\x_k\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix};\lambda^{T}e\le 1;\lambda\in\Re_{+}^{n}\right.\right\}.$$

Figure 3.2 displays the NIRS results for the example data set introduced in Section 1.3. In contrast to the FDH model, the NIRS model classifies DMU *B* as inefficient, because it is dominated by reference point *B'*, which is a linear combination of DMU *A* and DMU *C* (that represents a feasible input-output vector if the production technology satisfies NIRS).

.
.

---

[40] More specifically, the rate of convergence for the FDH model is $n^{-1/(p+m)}$, whereas the rate of convergence for the BCC model is $n^{-2/(p+m+1)}$. For more details, see Kneip *et al.* (1998) and Park *et al.* (1997).

[41] In the DEA literature, *returns-to-scale* refers to the marginal properties of the *production frontier*. By contrast, in economics, the term generally refers to the marginal properties of the *cost frontier* rather than the production frontier.

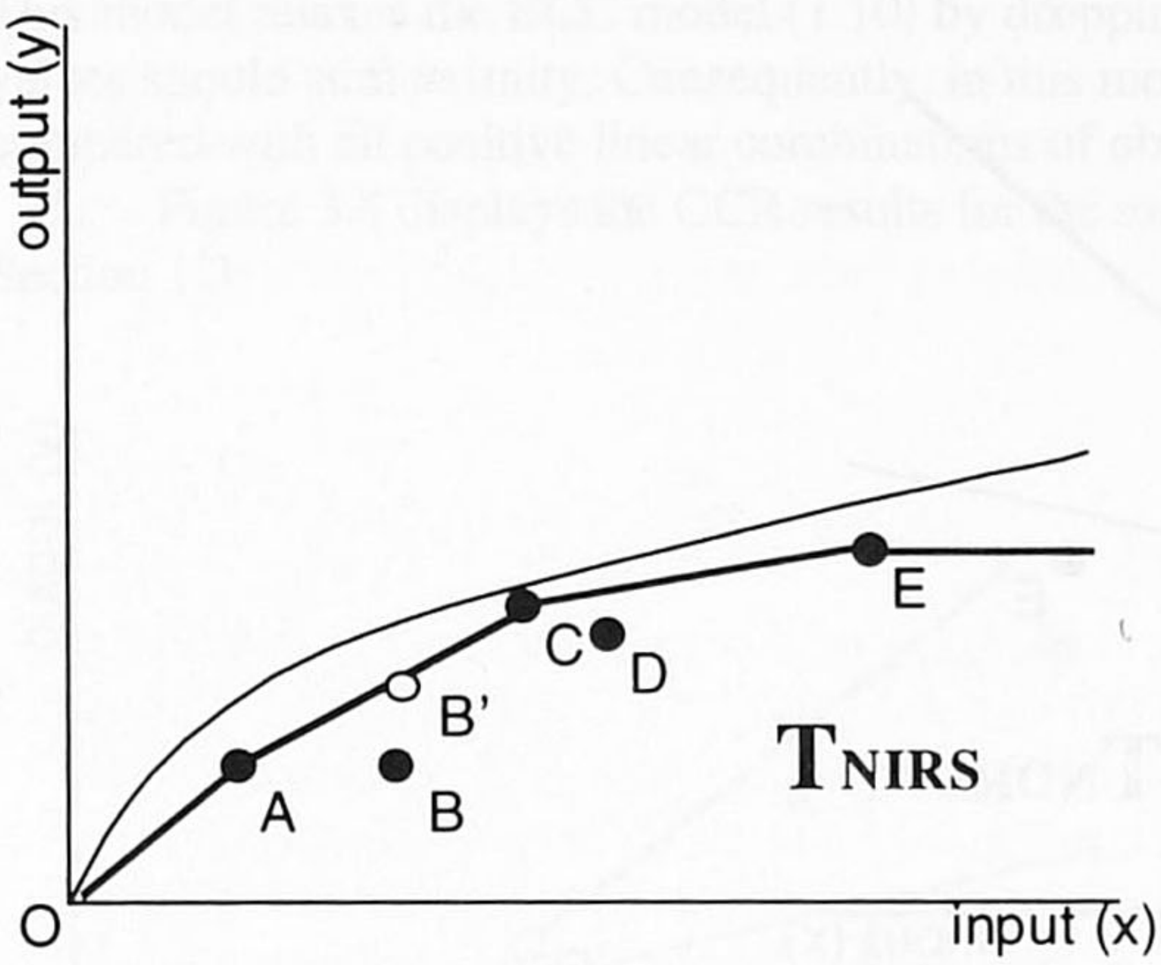


Figure 3. 2 The NIRS model for the example data set introduced in Section 1.3

This model is more progressive than the BCC model, because it allows for reference units that are constructed as positive linear combinations of observations (provided the sum of the weights does not exceed unity) in addition to convex combinations.

Similarly, Banker (1984) also proposed a model that incorporates the assumption that the production set satisfies *non-decreasing returns-to-scale* in addition to free disposability and convexity. This corresponds to assuming that, if a particular input-output vector is attainable, multiplying that vector with a scalar greater or equal than unity also gives an attainable vector, i.e.:

(3. 8) $(y,x)\in T \Rightarrow (ky,kx)\in T \quad \forall k\in[1,\infty)$.

The smallest subset in input-output space that is consistent with this assumption in addition to the BCC assumptions is the following empirical production set:

(3. 9) $$\hat{T}_{NDRS}=\left\{(y,x)\in\Re_+^{p+m}\left|\begin{pmatrix}y\\-x\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix};\lambda^T e\ge 1;\lambda\in\Re_+^n\right.\right\}.$$

Measuring efficiency relative to this set gives the following estimator:

(3. 10) $$\hat{\theta}_{NDRS,k}=\max_{\theta}\left\{\theta\left|(\theta y_k,x_k)\in\hat{T}_{NDRS}\right.\right\}$$

$$=\max_{\lambda,\theta}\left\{\theta\left|\begin{pmatrix}\theta y_k\\x_k\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix};\lambda^T e\ge 1;\lambda\in\Re_+^n\right.\right\}.$$

Figure 3.3 displays the NDRS results for the example data set introduced in Section 1.3.

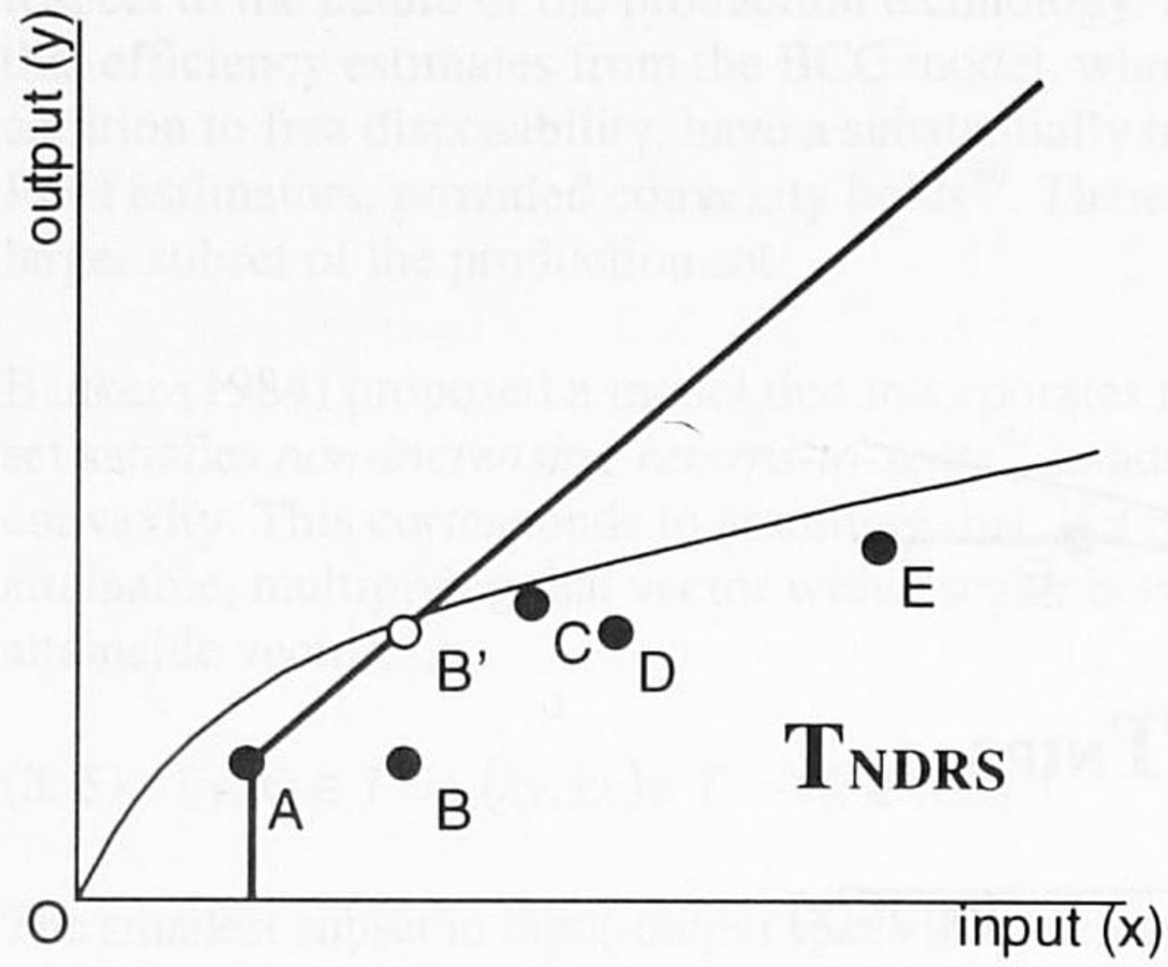


Figure 3. 3 The NDRS model for the example data set introduced

This model is more progressive than the BCC model, because it allows for reference units that are constructed as positive linear combinations of observations (provided the sum of the weights do not fall below unity) in addition to convex combinations. For example, the NIRS model compares DMU *B* with reference point *B'*, which is a multiple of DMU *A* that represents a more progressive input-output vector than the BCC reference point in Figure 1.1.

The original Charnes, Cooper and Rhodes (1978) model is even more restrictive than the NIRS and NDRS models. It incorporates the assumption that the production set satisfies *constant returns-to-scale* in addition to free disposability and convexity. This corresponds to assuming that, if a particular input-output vector is attainable, all positive multiples of that combination are also attainable, i.e.:

(3. 11) $(y,x)\in T \Rightarrow (ky,kx)\in T \quad \forall k\in(0,\infty)$.

If this assumption holds in addition to free disposability, convexity and distribution assumptions (1.5), (1.8) and (1.15), fictitious input-output vectors created as positive linear combinations of observations are attainable. Based on this insight, the CCR model measures efficiency relative to the following empirical production set:

(3. 12) $\hat{T}_{CCR}=\left\{(y,x)\in\Re_+^{p+m}\middle|\begin{pmatrix}y\\-x\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix},\lambda\in\Re_+^n\right\}$.

Measuring efficiency relative to this set gives the following estimator:

(3. 13) $\hat{\theta}_{CCR,k}=\max_{\theta}\left\{\theta\middle|(\theta y_k,x_k)\in\hat{T}_{CCR}\right\}$

$$=\max_{\lambda,\theta}\left\{\theta\middle|\begin{pmatrix}\theta y_k\\x_k\end{pmatrix}\le\begin{pmatrix}Y\lambda\\-X\lambda\end{pmatrix},\lambda\in\Re_+^n\right\}.$$

This model relaxes the BCC model (1.10) by dropping the restriction that the lambda values should sum to unity. Consequently, in this model the evaluated unit can be compared with all positive linear combinations of observations.

Figure 3.4 displays the CCR results for the example data set introduced in Section 1.3.

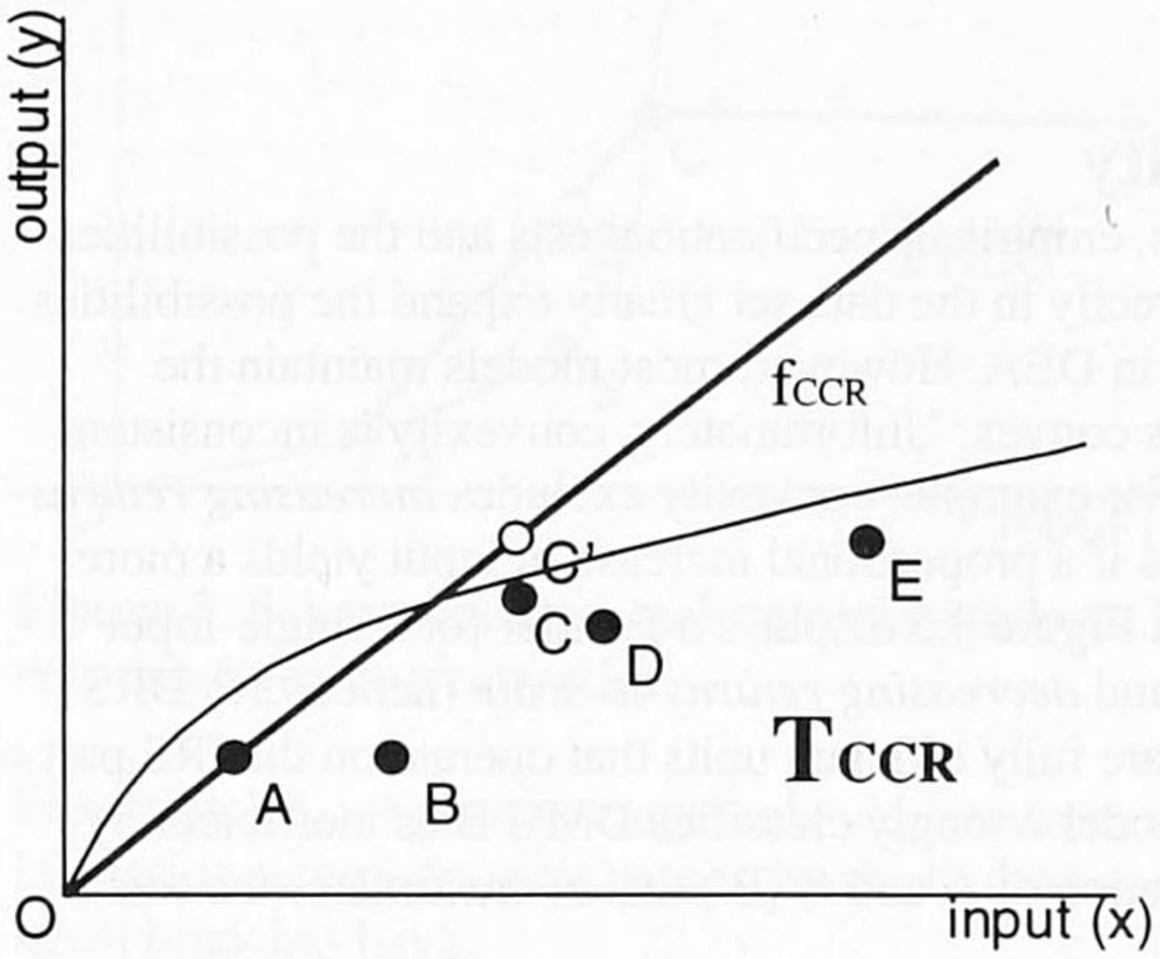


Figure 3. 4 The CCR model for the example data set introduced in Section 1.3

Like convexity, the assumptions with respect to the returns-to-scale properties can reduce finite sample error, because they yield a larger subset of the true production set, if the assumptions are satisfied. Unfortunately, imposing erroneous assumptions can cause *specification error*. For example, assuming constant returns-to-scale in case of variable returns-to-scale can systematically overestimate efficiency for units that do not operate at the most productive scale size. For example, DMU *C* in Figure 3.4 is approximately efficient. However, the CCR model wrongly indicates substantial possibilities for improvement, because DMU *C*'s production falls short of that of reference point *C*', which is a multiple of *A* that is beyond the true production possibilities. Increasing the sample can not reduce this source of error.

Consequently, DEA involves a difficult trade-off between finite sample error and specification error. It is generally desirable to select the most progressive set of correct production assumptions. Such a set gives the most efficient consistent estimator, i.e. the consistent estimator with less finite sample error than all other consistent estimators.

Specification tests have been developed, that allow for the empirical testing of the fit between the model and the data. For example, Banker (1993) proposed three asymptotic specification tests. Banker (1996) and Kittelsen (1997) evaluated the 'size' (i.e. true level of significance) and power of these tests using Monte-Carlo simulations. An interesting alternative test has been proposed by Pastor *et al.* (1995) and evaluated by Sirvent (1997).

The models above construct fictitious DMUs in different ways, e.g. convex combinations (BCC) and positive linear combinations (CCR). As discussed, these differences reflect differences in production assumptions, e.g. convexity (BCC) and constant returns-to-scale (CCR). Another way to introduce production information is to

introduce unobserved DMUs directly in the data set (Golany and Roll (1991) and Thannassoulis and Allen (1998)). The decision maker can create more than one unobserved DMU from a real DMU reflecting changes in marginal rates of substitution as one moves in some direction of changes to input-output levels set [42]. This allows for the inclusion of *local* production information in addition to *global* preference information.

3.3. Violations of convexity

The alternative model specifications, empirical specification tests and the possibilities for including unobserved DMUs directly in the data set greatly expand the possibilities for modelling production structures in DEA. However, most models maintain the assumption that the production set is convex. Unfortunately, convexity is inconsistent with certain production structures. For example, convexity excludes *increasing returns-to-scale* (henceforth IRS). IRS arises if a proportional increase in input yields a more than proportional increase in output. Figure 3.5 displays a frontier for a single-input single-output technology with IRS and *decreasing returns-to-scale* (henceforth DRS) characteristics. DMUs A, B, and C are fully efficient units that operate on the IRS part of the frontier[43]. However, the BCC model wrongly classifies DMU B as inefficient by comparing it with a convex combination of A and C (B') that overestimates the true production possibilities.

---

[42] Roll and Golany (1991) and Thannassoulis and Allen (1998) demonstrated that the introduction of unobserved DMUs is equivalent to imposing particular weight restrictions (see Chapter 2). However, weight restrictions affect the objective function, and are therefore linked to the preferences of the decision maker, whereas unobserved DMUs affect the admissable set and are therefore linked to the production possibilities.

[43] DMUs *A*, *B* and *C* are *technically* efficient, because the production set does not contain input-output vectors that involve more output and less input than these DMUs do (i.e. Pareto-Koopmans technical efficiency, see Chapter 1). However, the DMUs are not *economically* efficient, because the production set does contain input-output vectors that yield more profit at all positive prices for the inputs and outputs (see Chapter 2). Convexity can be imposed without harm for analysing economic efficiency, because the production set necessarily contains reference units with at least the profit associated with an arbitrary convex combination of observations (see Varian (1984)). However, this does not hold for technical efficiency.

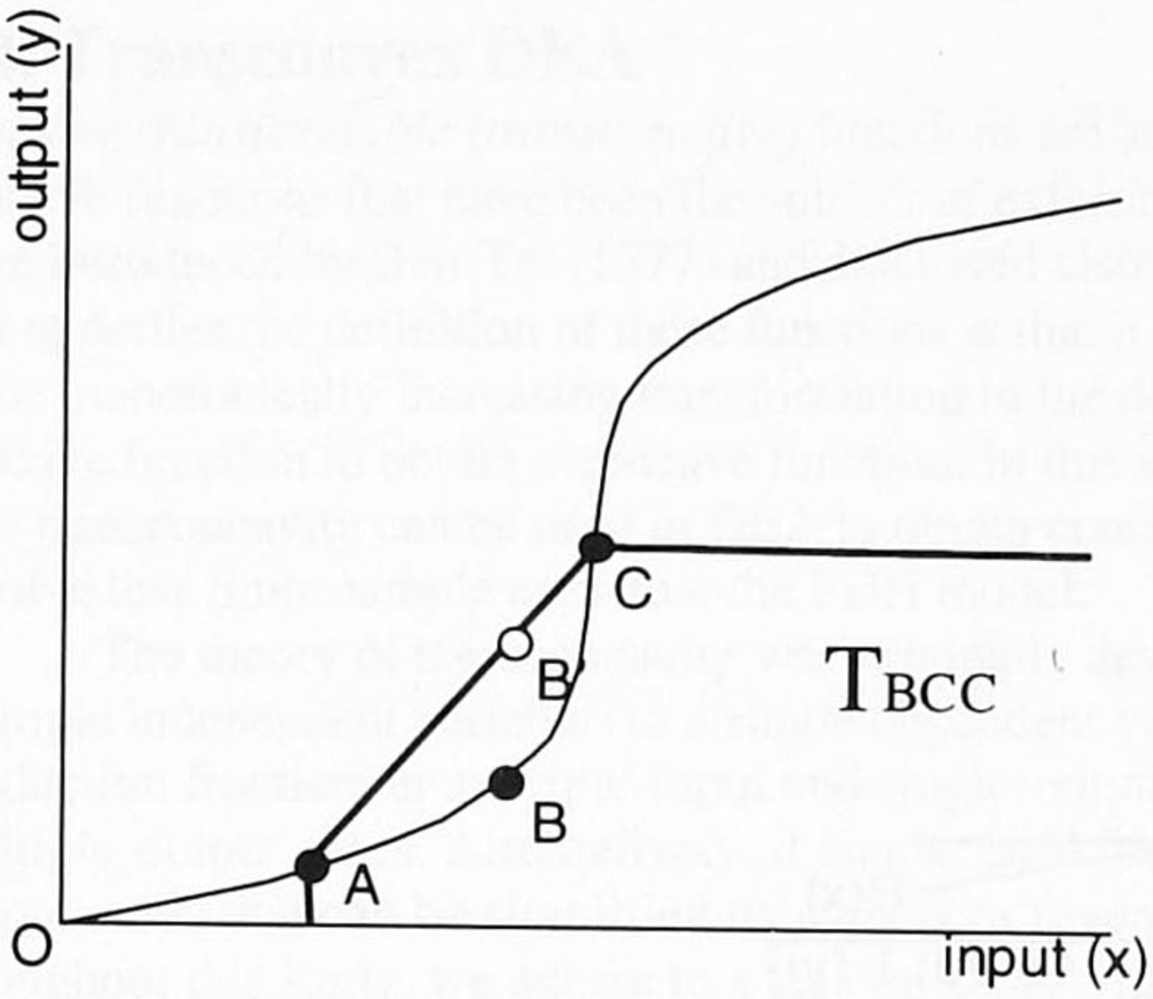


Figure 3. 5 A production technology with both increasing and decreasing returns-to-scale properties

In section 3.8, we present empirical evidence from a sample of bank branches for increasing returns-to-scale caused by the limited opportunities for division of labor that small branches have.

Increasing marginal rates of transformation (henceforth IMRT) for the outputs are another source of violation of convexity. IMRT arise if joint production of multiple output requires more input than separate production[44,45]. Figure 3.6 displays output isoquant $P(x)$ for a single-input two-output technology with IMRT. Both A, B and C are fully efficient and operate on the same isoquant. However, BCC compares B with a convex combination of A and C (B'), which substantially overestimates the true production possibilities.

---

[44] Increasing marginal rates of substitution are related to diseconomies of scope, which arises if joint production involves more costs than separate production (see Baumol, Panzar and Willig (1982)).

[45] Production sets are typically not assumed convex in microeconomics. However, the isoquants in input space (input sets) and the isoquants in output space (output sets) typically are assumed convex. This assumption excludes increasing marginal rates of substitution and increasing marginal rates of transformation. However, it does not exclude increasing returns-to-scale. Section 3.5 elaborates on the assumption of convex input and output sets.

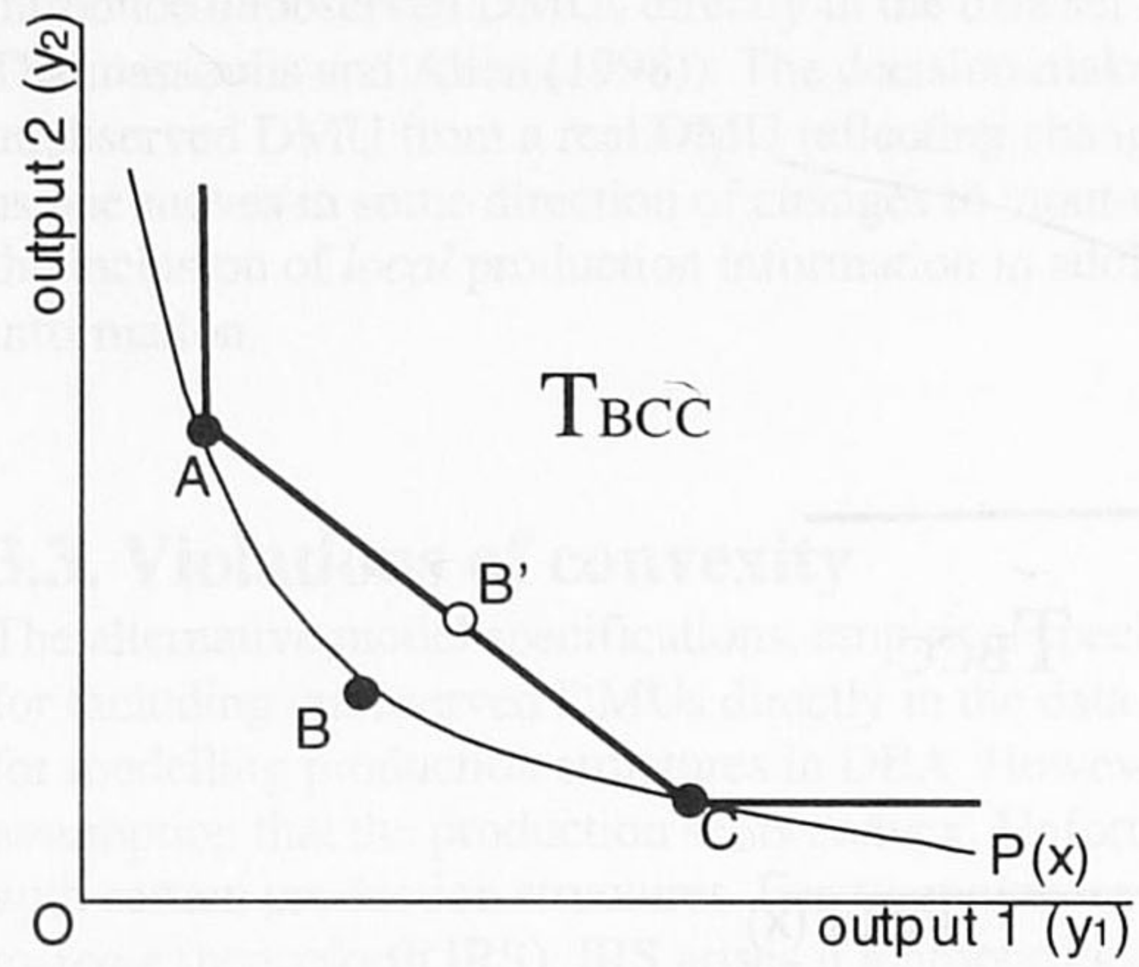


Figure 3. 6 A production technology with increasing marginal rates of transformation for the outputs

Another violation of convexity is *increasing marginal rates of substitution* (IMRS) or inputs that are complements instead of substitutes, i.e. inputs with a productivity that is higher if employed separately, rather than combined with other inputs. Figure 3.7 displays input isoquant *L(y)* for a technology with a single output and two inputs that are negative substitutes. Although A, B and C operate on the same isoquant, BCC wrongly classifies B as inefficient, because its production falls short of that of a convex combination of A and B (B') that exceeds the true production possibilities.

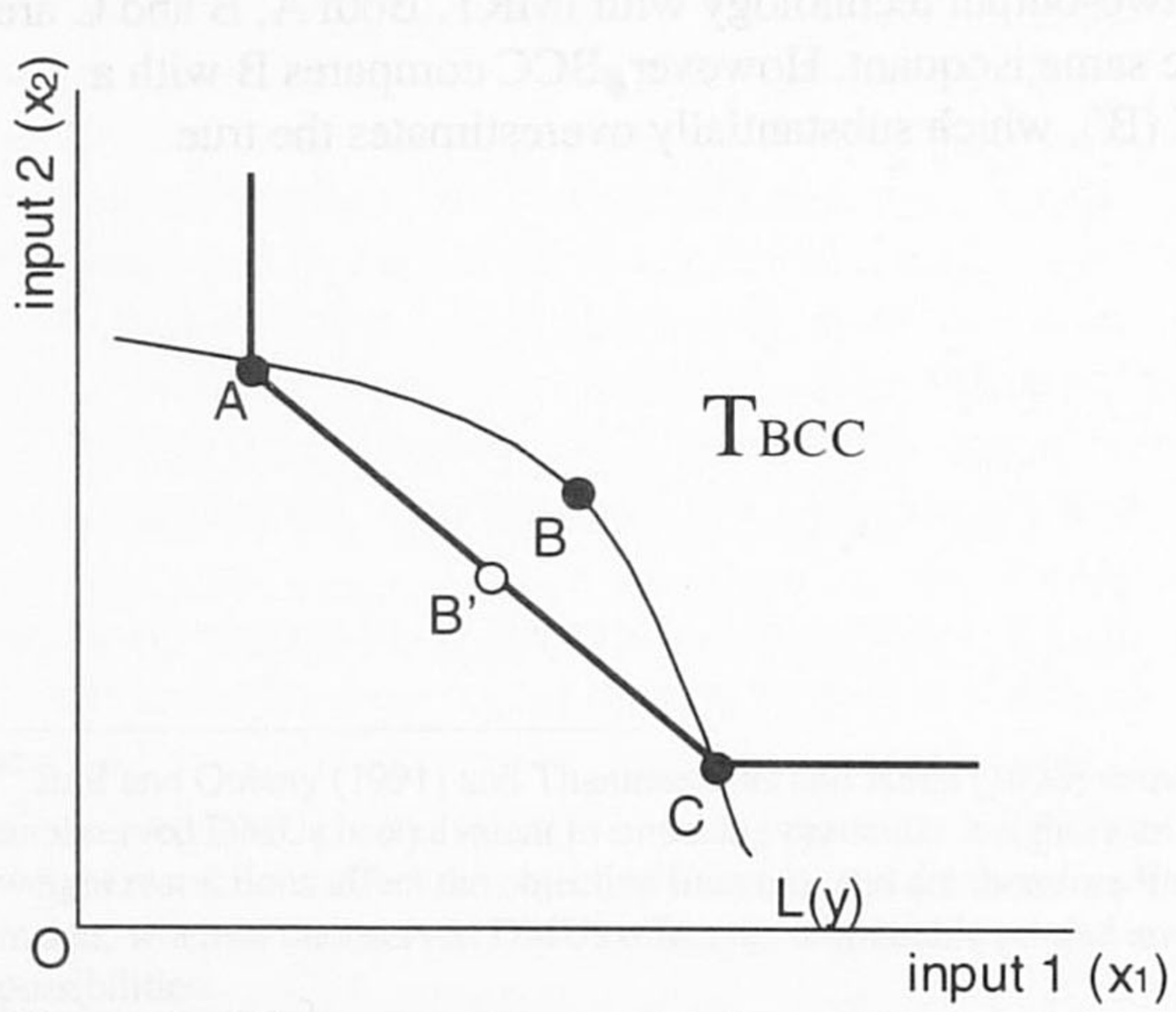


Figure 3. 7 A production technology with increasing marginal rates of substitution for the inputs

As discussed above, FDH can deal with such non-convexities, provided the production set is satisfies free disposability. However, given its finite sample performance, it is desirable to develop more efficient consistent non-convex models. In the next Section we will demonstrate how some key concepts from the mathematical theory of *Generalised Convexity* can help to achieve that goal.

3.4. Transconvex DEA

Concave transformable (*transconcave*) functions are an important class of generalised concave functions that have been the subject of extensive research. These functions were introduced by Ben-Tal (1977) and discussed also by Avriel *et al.* (1988). The idea that underlies the definition of these functions is that it is sometimes possible to apply some monotonically increasing transformation to the domain and/or range of a non-concave function to obtain a concave function. In this section we intend to demonstrate how transconcavity can be used in DEA to obtain consistent non-convex models that involve less finite sample error that the FDH model.

The theory of transconcavity was originally developed for functions that map multiple independent variables to a single dependent variable. Therefore, it applies to production frontiers in multiple-input and single-output cases and single-input and multiple-output cases. Alternatively, it can be used for multiple-input and multiple-output cases that can be simplified using cost or revenue functions. However, throughout this study, we adhere to a general DEA with multiple inputs and multiple inputs, for analysing technical efficiency. To preserve this level of generality, we generalise the mathematics of *transconcave functions* to a mathematics of *transconvex sets*. Specifically, we adhere to the following definition:

Definition 3.1 A production set T is *convex transformable* or *transconvex* if there exist transformation functions $G:\Re_+^p \to \Re_+^p$ (for outputs) and $H:\Re_+^m \to \Re_+^m$ (for inputs) such that the following transformed production set is convex:

(3. 14) $T' = \left\{(G(y),H(x)) \in \Re_+^{p+m} \middle| (y,x) \in T\right\}$.

Alternatively, transformations G and H convexify the production set, if the following condition holds:

(3. 15) $\left(G(Y)\lambda, H(X)\lambda\right) \in T' \quad \forall (Y,X) \in T; \lambda^T e = 1; \lambda \in \Re_+^n$. □

In this study, the convexifying transformations are assumed to be invertible[46], i.e.:

(3. 16) $\left(G^{-1}(G(y)), H^{-1}(H(x))\right) = (y,x) \quad \forall (y,x) \in T$.

If this condition holds, the following expression relates the original production set to the transformed production set:

(3. 17) $T = \left\{(G^{-1}(y), H^{-1}(x)) \in \Re_+^{p+m} \middle| (y,x) \in T'\right\}$.

Consequently, transconvexity condition (3.15) can be expressed in terms of the original production set in the following way:

(3. 18) $\left(G^{-1}(G(Y)\lambda), H^{-1}(H(X)\lambda)\right) \in T \quad \forall (Y,X) \in T; \lambda^T e = 1; \lambda \in \Re_+^n$.

---

[46] The invertibility assumption restricts the class of production sets that can be convexified. Unfortunately, we can not formally define the class of production sets that can be convexified using invertible transformations. However, an important class of production sets can be convexified using exponential transformations, which are invertible. Section 3.5 and 3.6 discuss these transformations in more detail.

This corresponds to the condition that particular *generalised mean values* (Hardy *et al.* 1952) of attainable input-output vectors are contained within the production set. More specifically, the following generalised mean value functions are employed:

Definition 3.2 The function $g(Y,\lambda):[0,1]\rightarrow\Re_+^p$ defined as

(3. 19) $g(Y,\lambda)=G^{-1}(G(Y)\lambda)$

represents the G-mean-value function for outputs. Similarly, the function $h(X,\lambda):[0,1]\rightarrow\Re_+^m$ defined as

(3. 20) $h(X,\lambda)=H^{-1}(H(Y)\lambda)$

represents the H-mean-value function for inputs. □

Before we show how this mathematics can help to build better non-convex DEA models, first note that input-output vectors that are Pareto-Koopmans dominated by generalised mean-values of observed input-output vectors are contained within the production set, if free disposability (1.6) holds in addition to transconvexity. This insight can be used to construct the following Transconvex Data Envelopment Analysis (TDEA) approximation for the production set:

$$\textbf{(3. 21)}\quad \hat{T}_{TDEA}^{G,H}=\left\{(y,x)\in\Re_+^{p+m}\left|\begin{pmatrix}y\\-x\end{pmatrix}\leq\begin{pmatrix}G^{-1}(G(Y)\lambda)\\-H^{-1}(H(X)\lambda)\end{pmatrix};\lambda^T e=1;\lambda\in\Re_+^n\right.\right\}.$$

Measuring efficiency relative to this set gives the following TDEA estimator:

$$\textbf{(3. 22)}\quad \hat{\theta}_{TDEA,k}^{G,H}=\max_{\theta}\left\{\theta\left|(\theta y,x)\in\hat{T}_{TDEA}^{G,H}\right.\right\}$$

$$=\max_{\lambda,\theta}\left\{\theta\left|G^{-1}(G(Y)\lambda)\geq\theta y_k;\ H^{-1}(H(X)\lambda)\leq x_k;\lambda^T e=1;\lambda\in\Re_+^n\right.\right\}.$$

The standard model takes convex combinations of the input-output vectors of the referencing DMUs. In formulation (3.22) these convex combinations are replaced with H-mean-values of inputs and G-mean-values of outputs.

If the production set satisfies transconvexity (3.18) in addition to free disposability (1.6), and if the observations satisfy homogeneity (1.5) and accuracy (1.8), the empirical production set is contained within the true production set, i.e. $\hat{T}_{TDEA}^{G,H}\subseteq T$. Therefore, the estimator is biased below true efficiency, i.e.:

$$\textbf{(3. 23)}\quad P\left(\hat{\theta}_{TDEA,k}^{G,H}-\theta_k>\varepsilon\right)=0\quad\forall\varepsilon>0.$$

However, increasing the sample size generally improves the estimation. More specifically, if the standard distribution assumption (1.15) holds in addition to the

above assumptions, the TDEA estimator is statistically consistent (i.e. asymptotically unbiased and with a vanishing variance), i.e.:

(3. 24) $\lim_{n\to\infty} P\left(\left|\hat{\theta}^{G,H}_{TDEA,k} - \theta_k\right| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0.$

Proof 3.1 If the data set contains a DMU that consumes at most x_k to produce at least $\theta_k y_k$, the estimator is necessarily greater than or equal to true efficiency, i.e.:

(I) $\exists j \in D : x_j \le x_k ; y_j \ge \theta_k y_k \Rightarrow \hat{\theta}^{G,H}_{TDEA,k} \ge \theta_k .$

In addition, $(\theta_k y_k, x_k) \in T$ holds by definition. Therefore, (1.15) implies

$$P(y_j \ge \theta_k y_k, x_j \le x_k) = F(\theta_k y_k, x) > 0 .$$

Consequently, the probability of observing at least one such DMU converges to unity as the sample increases, i.e.:

$$\lim_{n\to\infty} P\left(\exists j \in D : y_j \ge \theta_k y_k, x_j \le x_k\right) = \lim_{n\to\infty}\left(1 - (1 - F(\theta_k y_k, x_k))^n\right) = 1 .$$

Combining this with (I), we find that, asymptotically, the estimator is biased above true efficiency:

(II) $\lim_{n\to\infty} P\left(\hat{\theta}^{G,H}_{TDEA,k} - \theta_k < \varepsilon\right) = 0 \quad \forall \varepsilon > 0.$

Combining (II) and (3.23) gives consistency, i.e.:

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}^{G,H}_{TDEA,k} - \theta_k\right| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0. \square$$

The transconvex model allows for the introduction of a wide range of alternative production assumptions in the standard DEA framework. Any convex transformable production set can be estimated consistently by employing transformation functions that convexify that specific production set. However, it is desirable to select the weakest transformation that convexifies the production set. This is because strong transformations involve more finite sample error than weaker ones. Transformations G' and H' are stronger than transformations G and H if:

(3. 25) $\begin{pmatrix} G'^{-1}(G'(Y)\lambda) \\ -H'^{-1}(H'(X)\lambda) \end{pmatrix} \le \begin{pmatrix} G^{-1}(G(Y)\lambda) \\ -H^{-1}(H(X)\lambda) \end{pmatrix} \quad \lambda^T e = 1; \lambda \in \Re^n_+ .$

If G and H convexify T, we find

(3. 26) $\hat{T}^{G',H'}_{TDEA} \subseteq \hat{T}^{G,H}_{TDEA} \subseteq T .$

Therefore, stronger transformations involve more finite sample error than weaker ones:

(3. 27) $\left|\hat{\theta}^{G',H'}_{TDEA,k} - \theta_k\right| \geq \left|\hat{\theta}^{G,H}_{TDEA,k} - \theta_k\right|$.

Example 3.1 To illustrate the TDEA model, consider three DMUs (say *A*, *B* and *C*) that operate under a technology represented by the frontier $g(x) = 0.5x^2$. The input-output vectors are given by $(y_A, x_A) = (1,0.5)$, $(y_B, x_B) = (1.5,1.125)$, and $(y_C, x_C) = (2,2)$. Figure 3.8 displays the true frontier and the DEA frontier that is constructed from the observations.

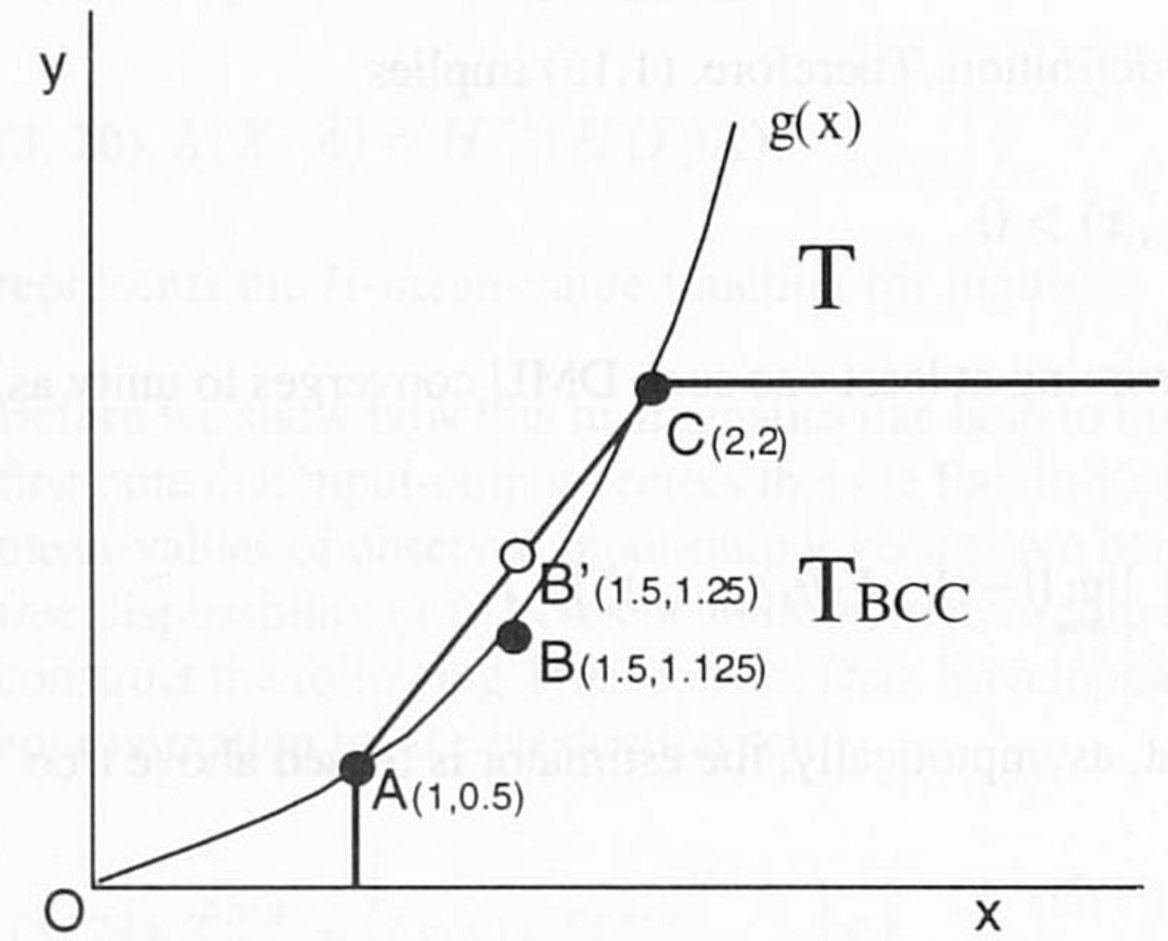


Figure 3. 8 Convex combinations in input-output space can represent unattainable input-output levels if the production set is non-convex

The true production set $T = \left\{(y,x) \in \Re^{p+m}_{+} \middle| y \leq 0.5x^2 \right\}$ is not convex. This implies that convex combinations of observations may be unattainable. For example, the standard DEA model compares DMU *B* with a convex combination of DMU *A* and DMU *C* (say *B'*) that is not attainable. However, the production set is transconvex, because transformations exist that convexify it. For example,
$G(y) = y$ and $H(x) = 0.5x^2$ convexify the production set. Figure 3.9 displays the true frontier and the DEA approximation in $(y, H(x) = 0.5x^2)$ space. Although the original production set is not convex, the transformed production set $T' = \left\{(y,0.5x^2) \in \Re^{p+m}_{+} \middle| y \leq 0.5x^2 \right\}$ is convex. DMU *B* is classified efficient, because no *G*- and *H*-mean-value of DMU *A* and DMU *C* dominates it.

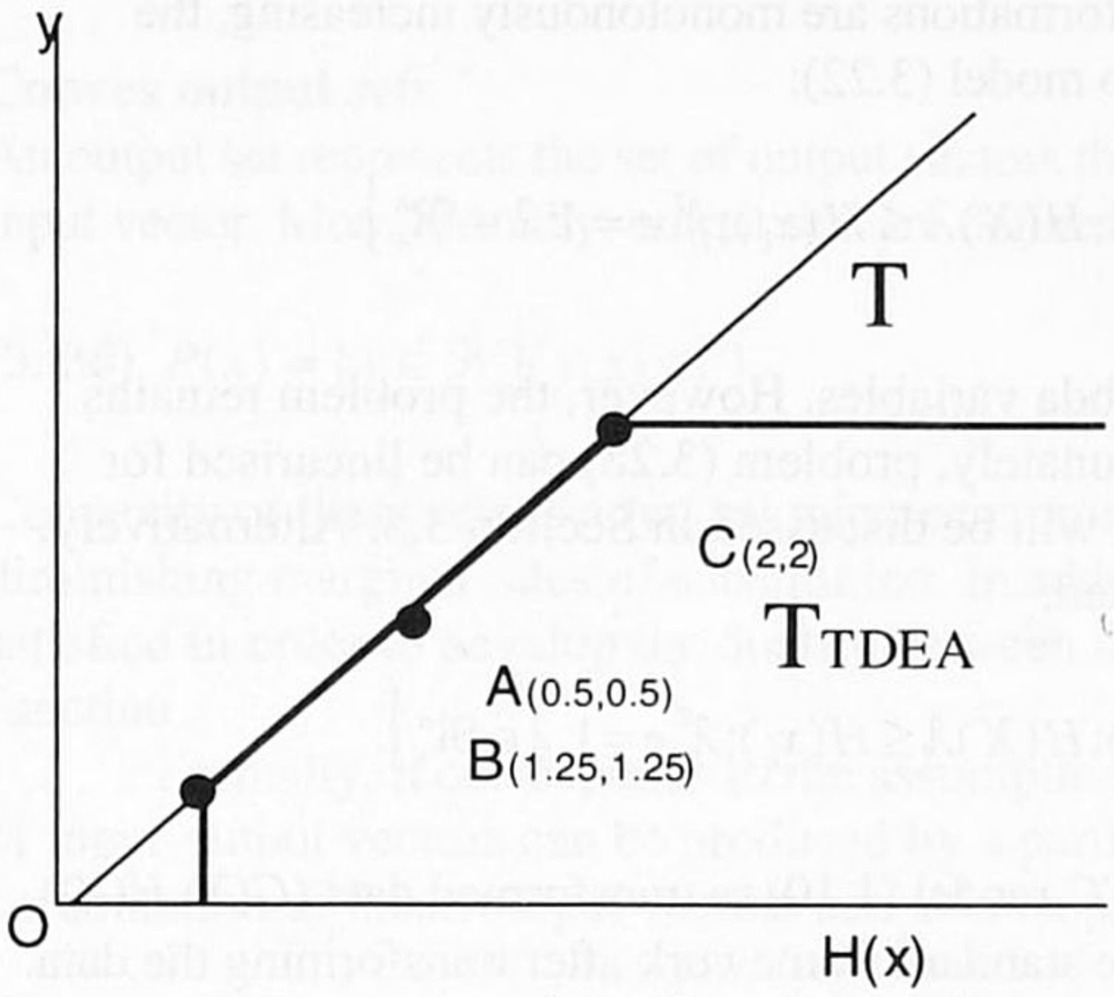


Figure 3. 9 Convex combinations in transformed input-output space do represent attainable input-output levels if the transformations convexify the production set

Similarly, $G(y) = y$ and $H'(x) = 0.25x^4$ convexify the production set, because the transformed production set $T' = \left\{(y, 0.25x^4) \in \Re_+^{p+m} \middle| y \le 0.5x^2 \right\}$ is convex. However, $H'(x) = 0.25x^4$ is stronger than $H(x) = 0.5x^2$ and hence shrinks the empirical production set. Therefore, $H'(x) = 0.25x^4$ involves more finite sample error than $H(x) = 0.5x^2$. Figure 3.10 displays the true frontier and the DEA frontier in $(y, H'(x))$ space.

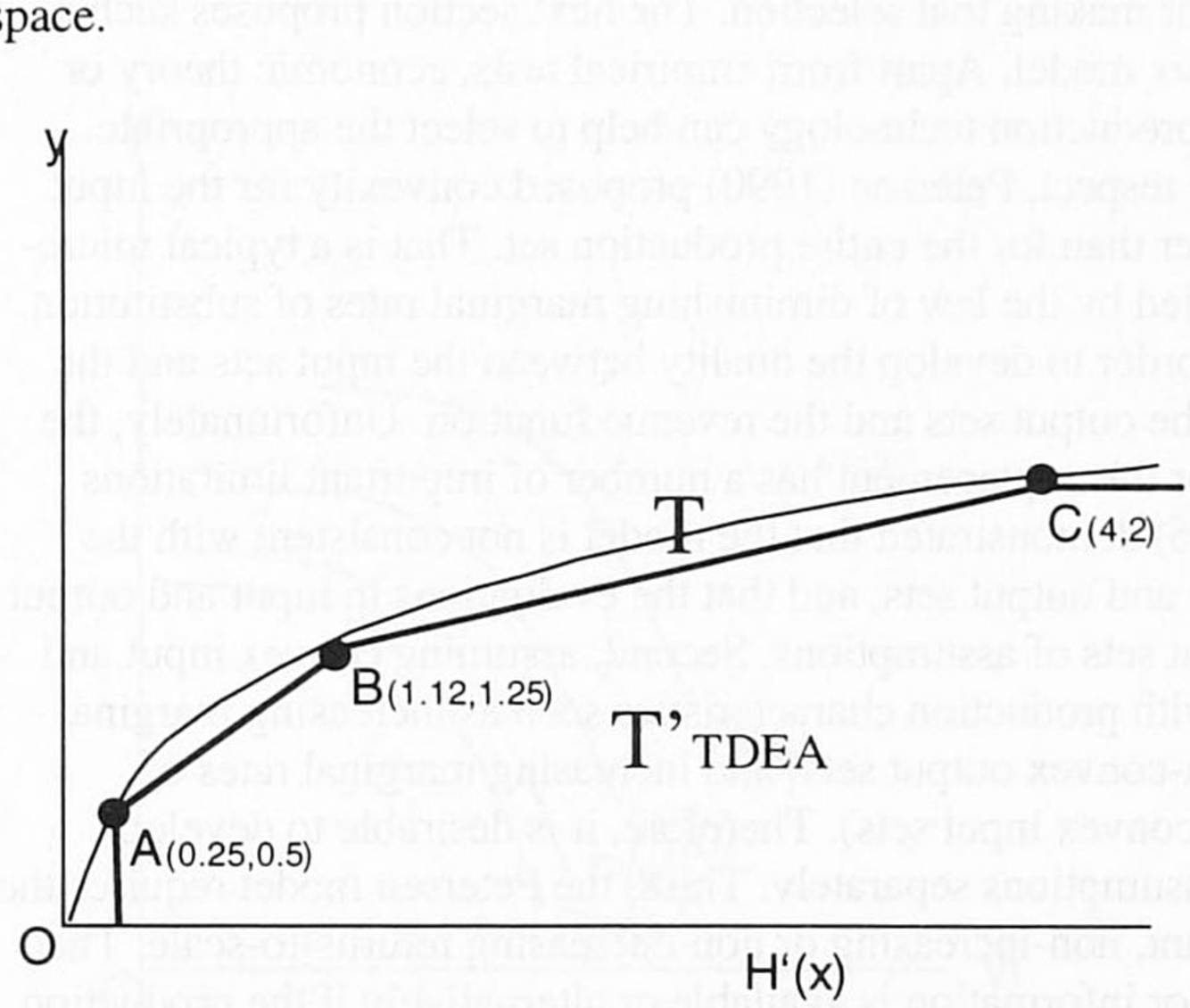


Figure 3. 10 A transformation that is stronger than required to convexify the production set can introduce finite sample error.

The TDEA estimator (3.22) can be non-linear. Although modern solvers should be capable of solving this non-linear formulation, it is desirable to find a linear formulation that can be solved using standard Linear Programming techniques.

If the domain and range transformations are monotonously increasing, the following formulation is equivalent to model (3.22):

$$\textbf{(3. 28)}\quad \hat{\theta}^{G,H}_{TDEA} = \max_{\lambda,\theta}\left\{\theta \middle| G(Y)\lambda \geq G(\theta y_k); H(X)\lambda \leq H(x_k); \lambda^T e = 1; \lambda \in \Re^n_+\right\}.$$

This formulation is linear in the lambda variables. However, the problem remains non-linear in the theta variable. Fortunately, problem (3.28) can be linearised for some important transformations that will be discussed in Section 3.5. Alternatively, the following linear model can be used:

$$\textbf{(3. 29)}\quad \hat{\theta}'^{G,H}_{TDEA} = \max_{\lambda}\left\{\theta \middle| G(Y)\lambda \geq \theta G(y_k); H(X)\lambda \leq H(x_k); \lambda^T e = 1; \lambda \in \Re^n_+\right\}.$$

This model essentially applies the BCC model (1.10) to transformed data ($G(Y)$,$H(X)$), and can therefore be solved within the standard framework after transforming the data. However, this model uses a non-radial projection to the TDEA frontier and does not give a consistent estimate for theta. Still, $\hat{\theta}'^{G,H}_{TDEA} = 1$ does give a necessary condition for efficiency.

3.5. Convex input sets and output sets

Different transformations correspond to different sets of production assumptions. As discussed above, the selection of the appropriate set of production assumptions involves a difficult trade-off between finite sample error and specification error. Empirical test procedures can be helpful for making that selection. The next section proposes such a procedure for the transconvex model. Apart from empirical tests, economic theory or practical knowledge of the production technology can help to select the appropriate model specification. In this respect, Petersen (1990) proposed convexity for the input sets and the output sets rather than for the entire production set. That is a typical micro-economic assumption justified by the law of diminishing marginal rates of substitution. In addition, it must hold in order to develop the duality between the input sets and the cost function and between the output sets and the revenue function. Unfortunately, the model Petersen proposed for this replacement has a number of important limitations.

First, Bogetoft (1996) demonstrated that the model is not consistent with the assumption of convex input and output sets, and that the evaluations in input and output spaces are based on different sets of assumptions. Second, assuming convex input and output sets is inconsistent with production characteristics such as increasing marginal rates of transformation (non-convex output sets) and increasing marginal rates of technical substitution (non-convex input sets). Therefore, it is desirable to develop models that impose these assumptions separately. Third, the Petersen model requires the prior specification of constant, non-increasing or non-decreasing returns-to-scale. That can be problematic if no prior information is available or alternatively if the production technology has both increasing and decreasing returns-to-scale properties. Finally, under non-increasing returns-to-scale, the model does not have a feasible solution if the data set contains observations that produce less of some output than the evaluated unit.

This Section introduces new models that impose convexity for input and/or output sets rather than convex production sets. The proposed models are embedded in the transconvex model discussed in the previous section. The models circumvent some of the important limitations inherent in the Petersen model.

Convex output sets

An output set represents the set of output vectors that can be produced from a particular input vector. More formally, output sets are defined as follows:

(3. 30) $P(x) = \{ y \in \Re_+^s \,|\, (y,x) \in T \}$.

Convexity of these sets is a typical microeconomic assumption, justified by the law of diminishing marginal rates of substitution. In addition, this assumption must be satisfied in order to develop the duality between the output set and the revenue function.

Formally, it corresponds to the assumption that, if all output vectors of a set F of input-output vectors can be produced by a particular input bundle, any convex combination of those output vectors can also be produced by that input bundle, i.e.:

(3. 31) $y_j \in P(x) \quad \forall j \in F \Rightarrow \sum_{j \in F} \lambda_j y_j \in P(x) \quad \sum_{j \in F} \lambda_j = 1; \lambda_j \geq 0 \; \forall j \in F$.

Figure 3.11 displays 4 output isoquants for a technology with two outputs and one input. The isoquants are progressively further away from eachother as the input level increases $(x_4 > x_3 > x_2 > x_1)$. The difference between the input levels is assumed constant, i.e. $x_4 - x_3 = x_3 - x_2 = x_2 - x_1$. This indicates increasing returns-to-scale, which violates convexity for the production set. However, each individual output set is convex.

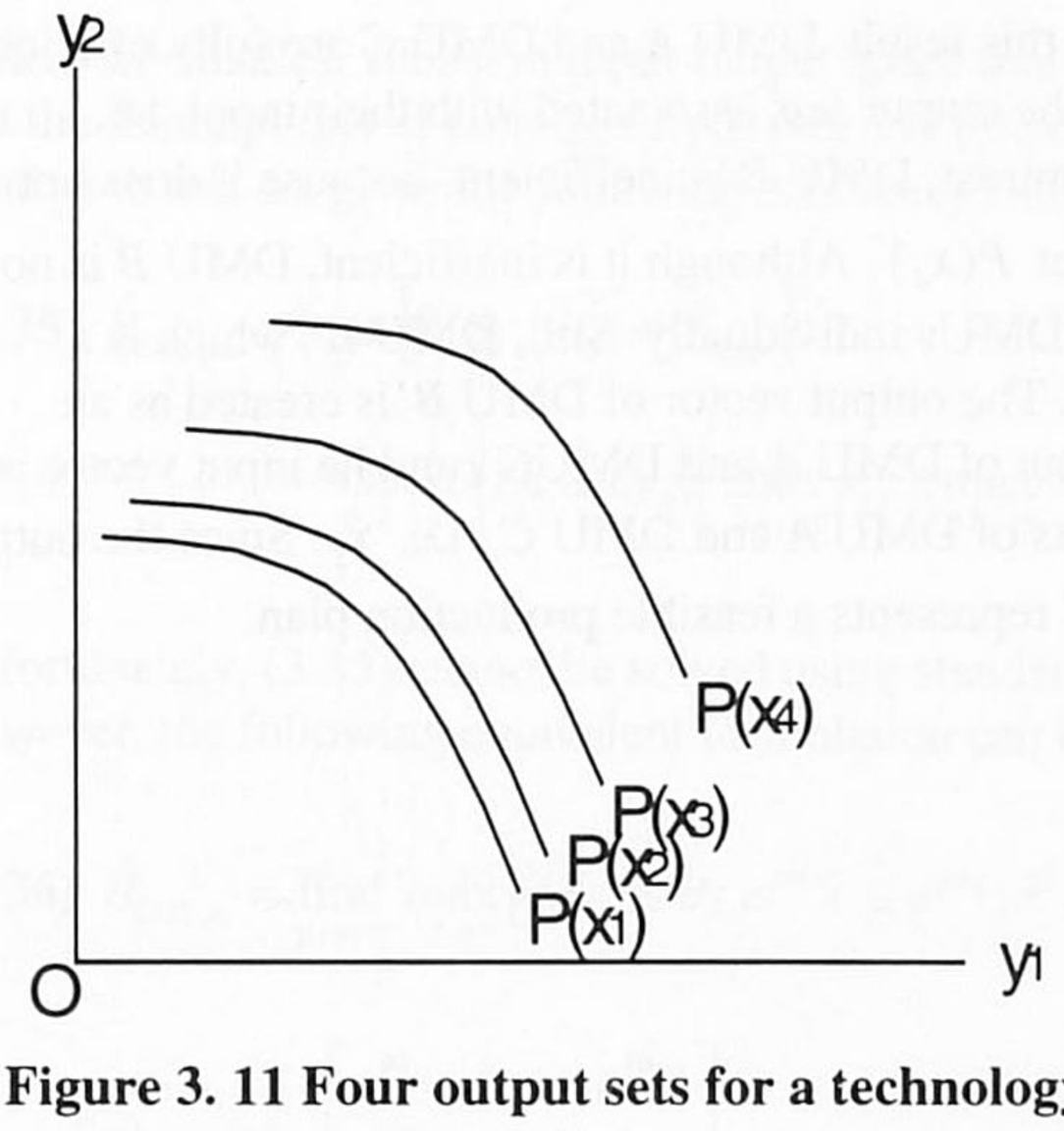


Figure 3. 11 Four output sets for a technology with increasing returns-to-scale properties

If the output sets are convex, and in addition homogeneity, free disposability and accuracy hold, any convex combination of observed output vectors could be produced

by the maximum of the observed input vectors of the units that constitute the convex combination, i.e.:

$$\textbf{(3. 32)}\quad Y\lambda \in P\left(\left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj}\right)\right) \quad \forall \lambda \in \Re_+^n : \lambda^T e = 1.$$

Proof 3. 2 If homogeneity (1.5), free disposability (1.6) and accuracy (1.8) hold, all observed output vectors of a subset of observations , say $F \subseteq D$, could be produced from the maximum of all produce output vectors of that subset, i.e.:

$$y_j \in P\left(\left(\max_{j\in F} x_{1j} \ldots \max_{j\in F} x_{mj}\right)\right) \quad \forall j \in F .$$

Therefore, if convexity for the input sets holds in addition to these assumptions, any convex combination of observed output vectors could also be produced by the maximum of the observed input vectors of the DMUs included in the convex combination, i.e. (3.32).□

This implies that the evaluated unit can harmlessly be compared with the following fictitious input-output vectors:

$$\textbf{(3. 33)}\quad \left\{\left(Y\lambda, \left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj}\right)\right) \middle| \lambda^T e = 1; \lambda \in \Re_+^n\right\}.$$

Example 3.2 Figure 3.12 illustrates this result. DMU A and DMU C are fully efficient, as they operate on the boundary of the output sets associated with their input, i.e. $P(x_1)$ and $P(x_3)$ respectively. By contrast, DMU B is inefficient, because it does not operate on the boundary of output set $P(x_4)$. Although it is inefficient, DMU B is not dominated by either of the efficient DMUs individually. Still, DMU B', which is a fictitious DMU, dominates DMU B. The output vector of DMU B' is created as an equally weighed average of the output of DMU A and DMUC , and he input vector is created as the maximum of the inputs of DMU A and DMU C , i.e. x_3 . Since the output sets are convex, this fictitious DMU represents a feasible production plan.

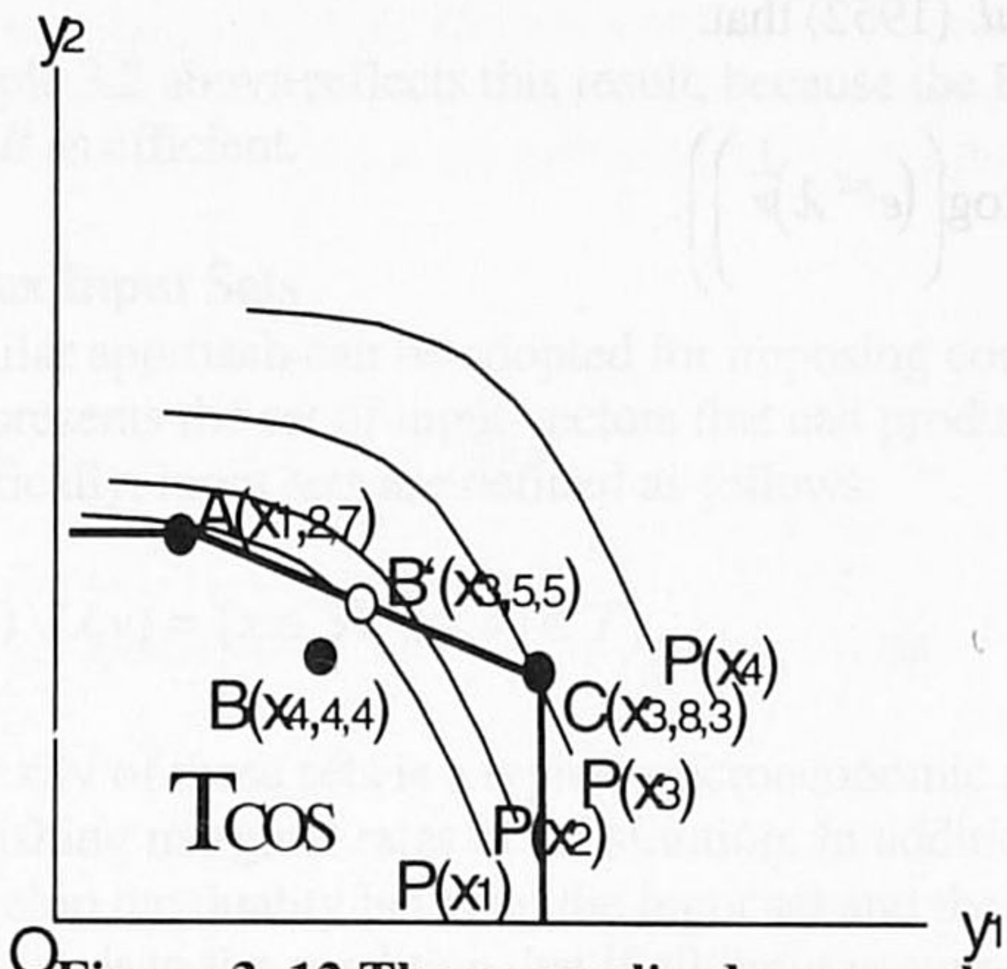


Figure 3. 12 The generalised mean value B' represents attainable input-output levels, because the output sets are convex.

This insight can be exploited to construct a model that takes advantage of the convexity of the output sets without imposing convexity for the complete production set. More specifically, we propose the following Convex Output Sets (COS) approximation for the production set:

$$\textbf{(3. 34)}\quad \hat{T}_{COS} = \left\{ (y,x) \in \Re_+^{p+m} \,\middle|\, y \le Y\lambda; x \ge \left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj} \right), \lambda^T e = 1; \lambda \in \Re_+^n \right\}.$$

This is the smallest subset in input-output space that is consistent with the observations and the assumptions of convex output sets and free disposability. Measuring efficiency relative to this set gives the following efficiency estimator:

$$\textbf{(3. 35)}\quad \hat{\theta}_{COS,k} = \max_{\theta} \left\{ \theta \,\middle|\, (\theta y_k, x_k) \in \hat{T}_{COS} \right\}$$

$$= \max_{\theta,\lambda} \left\{ \theta \,\middle|\, Y\lambda \ge \theta y_k; \left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj} \right) \le x_k; \lambda^T e = 1, \lambda \in \Re_+^n \right\}.$$

Unfortunately, (3.35) cannot be solved using standard Linear Programming techniques. However, the following equivalent formulation can be obtained:

$$\textbf{(3. 36)}\quad \hat{\theta}_{COS,k} = \lim_{p\to\infty} \left(\max_{\lambda,\theta} \left\{ \theta \,\middle|\, Y\lambda \le \theta y, e^{pX}\lambda \le e^{px_k}; \lambda^T e = 1; \lambda \in \Re_+^n \right\} \right).$$

Here e^{pX} denotes $\begin{bmatrix} e^{px_{11}} & \cdots & e^{px_{m1}} \\ \vdots & \ddots & \vdots \\ e^{px_{1n}} & \cdots & e^{px_{mn}} \end{bmatrix}$ for simplicity, and $e^{px_k} = \begin{bmatrix} e^{px_{1k}} \\ \vdots \\ e^{px_{mk}} \end{bmatrix}$.

Proof 3. 3 We know from Hardy *et al.* (1952) that:

$$\textbf{(I)} \qquad \left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj} \right) = \lim_{p\to\infty} \left(\log\left(\left(e^{pX} \lambda \right)^{\frac{1}{p}} \right) \right).$$

Therefore, restriction

$$\textbf{(II)} \qquad \left(\max_{j:\lambda_j>0} x_{1j} \ldots \max_{j:\lambda_j>0} x_{mj} \right) \leq x_k$$

is equivalent to:

$$\textbf{(III)} \qquad \lim_{p\to\infty} \left(\log\left(\left(e^{pX} \lambda \right)^{\frac{1}{p}} \right) \right) \leq x_k \, .$$

Since $\frac{1}{p}\ln(\cdot)$ with $p > 0$ is a non-decreasing transformation, this condition is equivalent to:

$$\textbf{(IV)} \qquad \lim_{p\to\infty} \left(e^{pX} \lambda \right) \leq \lim_{p\to\infty} \left(e^{px_k} \right).$$

Substituting (IV) in 3.35 gives 3.36.□

Formulation (3.36) can be approximated by the following Linear Programming formulation:

$$\textbf{(3. 37)} \quad \hat{\theta}_{COS,k} \approx \max_{\lambda,\theta} \left\{ \theta \middle| Y\lambda \leq \theta y; e^{pX} \lambda \leq e^{px_k}; \lambda^T e = 1; \lambda \in \Re_+^n \right\},$$

with p selected large enough.

This estimator is a special case of TDEA model (3.28). $G(y) = y$ and $H(x) = \lim_{p\to\infty} e^{px}$. It follows from (3.23) that the estimator is biased below true efficiency, i.e.:

$$\textbf{(3. 38)} \quad P\left(\hat{\theta}_{COS,k} - \theta_k > \varepsilon \right) = 0 \quad \forall \varepsilon > 0 \, .$$

However, increasing the sample size generally improves the estimator. More specifically, it follows from (3.24) that the estimator is statistically consistent, i.e.:

$$\textbf{(3. 39)} \quad \lim_{n\to\infty} P\left(\left| \hat{\theta}_{COS,k} - \theta_k \right| > \varepsilon \right) = 0 \quad \forall \varepsilon > 0.$$

In addition, if the output sets are convex, property (3.27) implies that the COS estimator has less finite sample error than FDH, i.e.:

$$\textbf{(3. 40)} \quad \left| \hat{\theta}_{COS,k} - \theta_k \right| \leq \left| \hat{\theta}_{FDH,k} - \theta_k \right| .$$

Example 3.2 above reflects this result, because the FDH model wrongly classifies DMU*B* as efficient.

Convex Input Sets

A similar approach can be adopted for imposing convexity for the input sets. An input set represents the set of input vectors that can produce a particular output vector. Specifically, input sets are defined as follows:

(3. 41) $L(y) = \{x \in \Re_+^m \,|\, (y,x) \in T\}$.

Convexity of these sets is a typical microeconomic assumption, justified by the law of diminishing marginal rates of substitution. In addition, this property is required in order to develop the duality between the input set and the cost function. Formally, it corresponds to the condition that if all input vectors of a set F of input-output vectors can produce at least a particular output bundle, all convex combinations of those input vectors can also produce that output bundle, i.e.:

(3. 42) $x_j \in L(y) \quad \forall j \in F \Rightarrow \sum_{j \in F} \lambda_j x_j \in L(y) \quad \sum_{j \in F} \lambda_j = 1; \lambda_j \geq 0 \; \forall j \in F$.

Figure 3.13 displays 4 input isoquants for a technology with one output and two inputs. The isoquants are progressively closer to each other as the output level increases $(y_4 > y_3 > y_2 > y_1)$. The difference between the output levels is assumed constant, i.e. $y_4 - y_3 = y_3 - y_2 = y_2 - y_1$. This indicates increasing returns-to-scale, which violates convexity for the production set. However, each individual input set is convex.

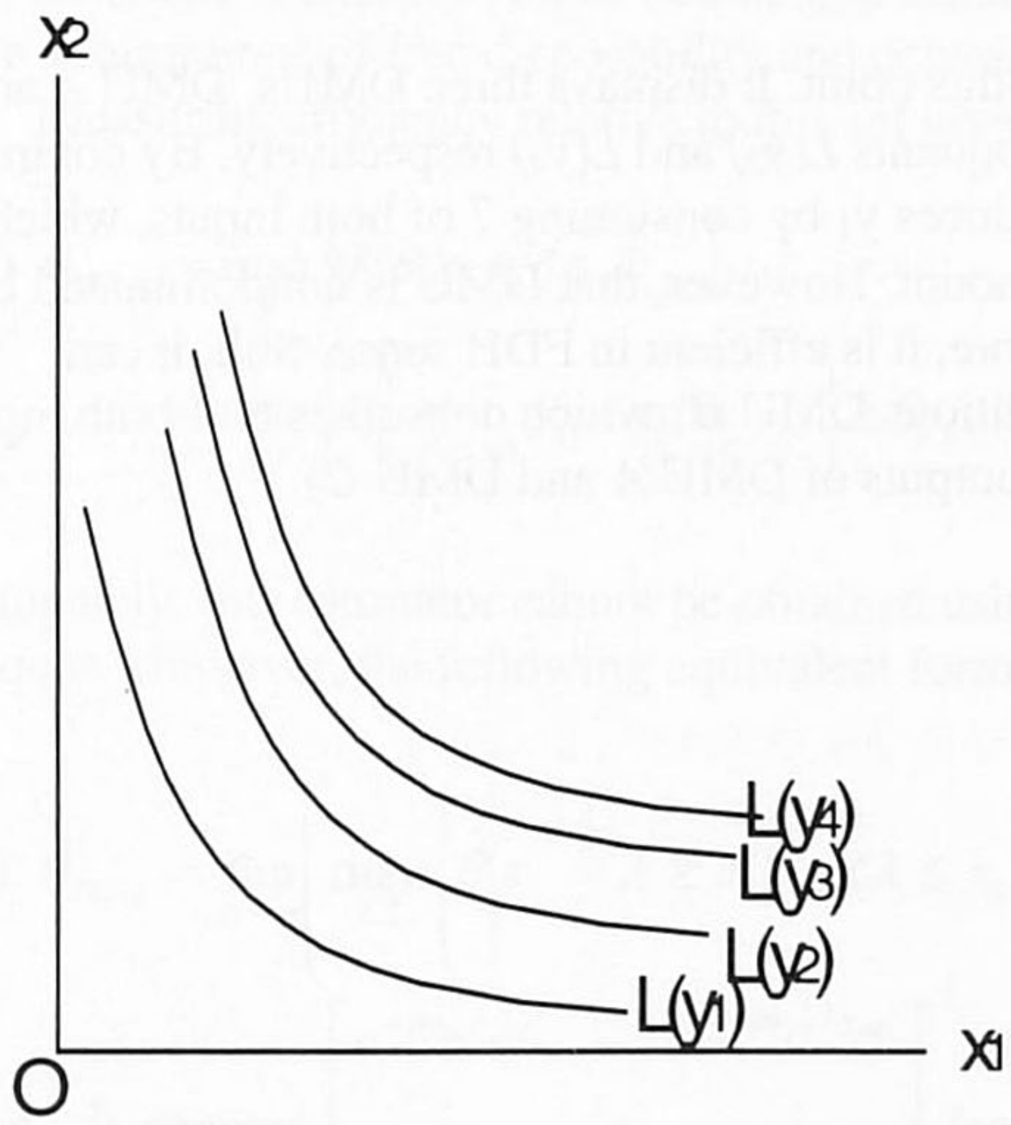


Figure 3. 13 Four input sets for a technology with increasing returns-to-scale properties

If convexity for the input sets holds in addition to homogeneity (1.5), free disposability (1.6) and accuracy (1.8), any convex combination of observed inputs could produce at least the minimum of the observed outputs of the units that constitute the convex combination, i.e.:

$$\textbf{(3. 43)}\quad X\lambda \in L\left(\left(\min_{j:\lambda_j>0} y_{1j} \ldots \min_{j:\lambda_j>0} y_{pj}\right)\right) \quad \lambda^T e = 1; \lambda \in \Re_+^n .$$

Proof 3. 4 If homogeneity (1.4), free disposability (1.5) and accuracy (1.7) hold, all observed input vectors of a subset of observations, say $F \subseteq D$, could produce at least the minimum of all output vectors, i.e.:

$$x_j \in L\left(\left(\min_{j\in F} y_{1j} \ldots \min_{j\in F} y_{pj}\right)\right) \quad \forall j \in F .$$

Therefore, if convexity for the input sets holds in addition to these assumptions, any convex combination of observed input vectors could also produce at least the minimum of the observed output vectors of the DMUs included in the convex combination, i.e. (3.43).□

This implies that the evaluated unit can harmlessly be compared with all fictitious input-output vectors

$$\textbf{(3. 44)}\quad \left\{\left(\left(\min_{j:\lambda_j>0} y_{1j} \ldots \min_{j:\lambda_j>0} y_{sj}\right), X\lambda\right) \middle| \lambda^T e = 1; \lambda \in \Re_+^n \right\}$$

Example 3.3 Figure 3. 14 illustrates this point. It displays three DMUs. DMU *A* and DMU *C* are efficient, operating at isoquants $L(y_2)$ and $L(y_4)$ respectively. By contrast, DMU *B* is inefficient, because it produces y_1 by consuming 7 of both inputs, which substantially exceeds the efficient amount. However, that DMU is not dominated by either of the efficient DMUs. Therefore, it is efficient in FDH sense. Still, it can harmlessly be compared with the fictitious DMU *B'*, which consumes 6 of both input to produce y_2 (i.e. the minimum of the outputs of DMU *A* and DMU *C*).

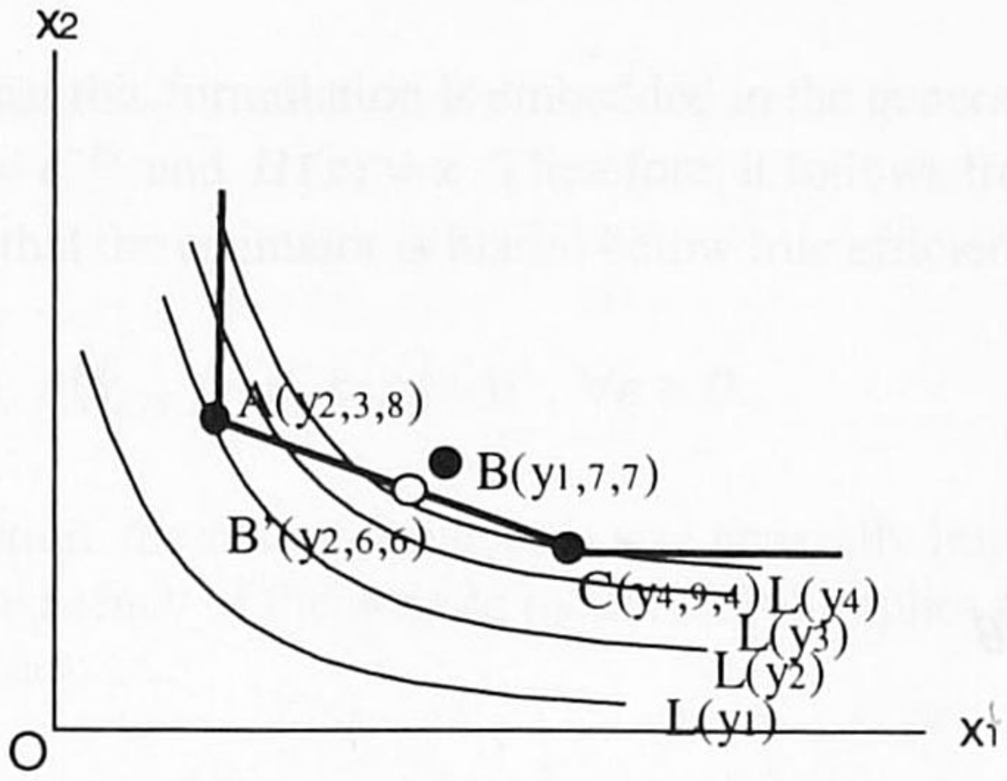


Figure 3. 14 The generalised mean value *B'* represents attainable input-output levels, because the input sets are convex.

This insight can be used to construct an empirical production set that takes advantage of the convexity for the input set. We propose the following Convex Input Sets (CIS) approximation for the production set:

$$\textbf{(3. 45)}\ \hat{T}_{CIS} = \left\{ (y,x) \in \Re_+^{p+m} \middle| y \le \left(\min_{j:\lambda_j>0} y_{1j} \ldots \min_{j:\lambda_j>0} y_{sj} \right); x \ge X\lambda; \lambda^T e = 1; \lambda \in \Re_+^n \right\}.$$

This is the smallest subset in input-output space that is consistent with the observations and the assumptions of free disposability and convex input sets.

Measuring efficiency relative to this set gives the following efficiency estimator:

$$\textbf{(3. 46)}\ \hat{\theta}_{CIS,k} = \max_{\theta} \left\{ \theta \middle| (\theta y_k, x_k) \in \hat{T}_{CIS} \right\}$$

$$= \max_{\lambda,\theta} \left\{ \theta \middle| \left(\min_{j:\lambda_j>0} y_{1j} \ldots \min_{j:\lambda_j>0} y_{sj} \right) \ge \theta y_k; X\lambda \le \theta x_k; \lambda^T e = 1; \lambda \in \Re_+^n \right\}.$$

Unfortunately, this estimator cannot be obtained using standard Linear Programming techniques. However, the following equivalent formulation can be obtained:

$$\textbf{(3. 47)}\ \hat{\theta}_{CIS,k} = \lim_{p\to\infty} \left(\max_{\lambda,\xi} \left\{ \theta \middle| e^{-p\frac{Y}{y_k}} \lambda \ge e^{-p\theta}; X\lambda \le x_k; \lambda^T e = 1; \lambda \in \Re_+^n \right\} \right).$$

Here, $e^{-p\frac{Y}{y_k}}$ denotes $\begin{bmatrix} e^{-py_{11}/y_{1k}} & \cdots & e^{-py_{p1}/y_{pk}} \\ \vdots & \ddots & \vdots \\ e^{-py_{1n}/y_{1k}} & \cdots & e^{-py_{pn}/y_{pk}} \end{bmatrix}$ for simplicity.

Proof 3. 5 The condition

$$\textbf{(I)} \qquad \left(\min_{j:\lambda_j>0} y_{1j} \ldots \min_{j:\lambda_j>0} y_{pj} \right) \geq \theta y_k$$

is equivalent to:

$$\textbf{(II)} \qquad \left(\min_{j:\lambda_j>0} \left(\frac{y_{ij}}{y_{1k}} \right) \ldots \min_{j:\lambda_j>0} \left(\frac{y_{pj}}{y_{pk}} \right) \right) \geq \theta \, .$$

In addition, we know from Hardy *et al.* (1952) that:

$$\textbf{(III)} \qquad \left(\min_{j:\lambda_j>0} \left(\frac{y_{ij}}{y_{1k}} \right) \ldots \min_{j:\lambda_j>0} \left(\frac{y_{pj}}{y_{pk}} \right) \right) = \lim_{p\to\infty} \left(\log \left(\left(e^{-p\frac{Y}{y_k}} \lambda \right)^{-\frac{1}{p}} \right) \right).$$

Consequently, condition (II) is equivalent to

$$\textbf{(IV)} \qquad \lim_{p\to\infty} \left(\log \left(\left(e^{-p\frac{Y}{y_k}} \lambda \right)^{-\frac{1}{p}} \right) \right) \geq \theta \, .$$

Since $\log(\cdot)^{-\frac{1}{p}}$ with $p > 0$ is a non-increasing transformation, the latter condition is equivalent to

$$\textbf{(V)} \qquad \lim_{p\to\infty} \left(e^{-p\frac{Y}{y_k}} \lambda \right) \leq \lim_{p\to\infty} \left(e^{-p\theta} \right).$$

Substituting this condition for (I) in the CIS estimator (3.46) gives the equivalent formulation (3.47).□

This formulation can be approximated by the following Linear Programming formulation:

$$\textbf{(3. 48)} \quad e^{-p\hat{\theta}_{CIS,k}} \approx \min_{\lambda,\xi} \left\{ \xi \left| e^{-p\frac{Y}{y_k}} \lambda \geq \xi; X\lambda \leq x_k; \lambda^T e = 1; \lambda \in \Re_+^n \right. \right\},$$

with p selected large enough.

Note that this formulation is embedded in the general TDEA model (3.28), with $G(y) = e^{-py}$ and $H(x) = x$. Therefore, it follows from the bias of the general model (3.23) that the estimator is biased below true efficiency:

(3. 49) $P\left(\hat{\theta}_{CIS,k} - \theta_k > \varepsilon\right) = 0 \quad \forall \varepsilon > 0$.

In addition, increasing the sample size generally improves the estimation. Specifically, the consistency of the general model (3.24) implies that the CIS model is statistically consistent, i.e.:

(3. 50) $\lim_{n\to\infty} P\left(\left|\hat{\theta}_{CIS,k} - \theta_k > \varepsilon\right|\right) = 0 \quad \forall \varepsilon > 0$.

In addition, if the output sets are convex, property (3.27) implies that the CIS estimator has less finite sample error than FDH, i.e.:

(3. 51) $\left|\hat{\theta}_{CIS,k} - \theta_k\right| \leq \left|\hat{\theta}_{FDH,k} - \theta_k\right|$.

Example 3.3 above reflects this result, because FDH wrongly classifies DMU *B* as efficient.

Some important extensions

The CIS and COS models impose either convexity for the input sets or convexity for the output sets. Since both assumptions are necessary to develop the duality between the input set and the cost function and between the output set and the revenue function, it is useful to develop a model that imposes both. For that purpose, we propose to use the maximum of the CIS and COS estimators, i.e.:

(3. 52) $\hat{\theta}_{CIOS,k} = \max\left(\hat{\theta}_{CIS,k}, \hat{\theta}_{COS,k}\right)$.

Since $\hat{\theta}_{CIS,k}$ and $\hat{\theta}_{COS,k}$ are both consistent (3.39 and 3.50), this model gives statistically consistent estimators for all technologies with monotone non-decreasing and convex input and output sets, provided the standard distribution assumptions hold, i.e.:

(3. 53) $\lim_{n\to\infty} P\left(\left|\hat{\theta}_{CIOS,k} - \theta_k > \varepsilon\right|\right) = 0 \quad \forall \varepsilon > 0$.

In addition, since both models are biased below true efficiency (3.40 and 3.51), CIOS involves less finite sample error than either does individually, i.e.:

(3. 54) $\left|\hat{\theta}_{CIOS,k} - \theta_k\right| \leq \left|\hat{\theta}_{CIS,k} - \theta_k\right|$,

(3. 55) $\left|\hat{\theta}_{CIOS,k} - \theta_k\right| \leq \left|\hat{\theta}_{COS,k} - \theta_k\right|$.

The above CIS, COS and CIOS models do not impose assumptions with respect to the returns-to-scale properties of the production set. However, if information regarding

the returns-to-scale properties is available, it can be used to help improve finite sample performance. For simplicity, we focus on including such information in the CIS model (3.5). Extensions for COS and CIOS are analogous.

Returns-to-scale assumptions can be represented by the general assumption that, if a particular input-output vector is technically possible, it also possible to produce that combination multiplied by a scale factor k, i.e.:

(3. 56) $(x,y)\in T \Rightarrow \quad (kx,ky)\in T \quad k\in \Gamma$.

$\Gamma=[0,\infty>$ corresponds to constant returns-to-scale, $\Gamma=[0,1]$ to non-increasing returns-to-scale, $\Gamma=[1,\infty>$ to non-decreasing returns-to-scale, and $\Gamma=1$ to variable returns-to-scale.

If this condition holds, in addition to the standard assumptions of homogeneity (1.5), free disposability (1.6) and accuracy (1.8), and convex input sets (3.42), all linear combinations of observed input bundles with a scale factor (i.e. $\lambda^T e$) that falls in Γ can produce at least the minimum of the observed outputs multiplied by that scale factor, i.e.:

$$\textbf{(3. 57)}\quad \lambda^T X \in L\left(\left(\min_{j:\lambda_j>0} y_{1j}\ldots \min_{j:\lambda_j>0} y_{sj}\right)\lambda^T e\right) \quad \lambda\in\Re_+^n : \lambda^T e\in\Gamma .$$

This implies that the evaluated unit can harmlessly be compared with all fictitious input-output vectors

$$\textbf{(3. 58)}\quad \left\{\left(\left(\min_{j:\lambda_j>0} y_{1j}\ldots \min_{j:\lambda_j>0} y_{sj}\right), X\lambda\right)\middle|\lambda^T e\in\Gamma;\lambda\in\Re_+^n\right\}.$$

This insight can be exploited by using the following approximation for the production set:

$$\textbf{(3. 59)}\quad \hat{T}'_{CIS}=\left\{(y,x)\in\Re_+^{p+m}\middle| y\geq\left(\min_{j:\lambda_j>0} y_{1j}\ldots \min_{j:\lambda_j>0} y_{sj}\right), x\leq X\lambda;\lambda^T e\in\Gamma;\lambda\in\Re_+^n\right\}.$$

Measuring efficiency relative to this set gives the following efficiency estimator:

$$\textbf{(3. 60)}\quad \hat{\theta}'_{CIS,k}=\min_{\theta}\left\{\theta\middle|(\theta x_k,y_k)\in\hat{T}'_{CIS}\right\}$$
$$=\max_{\lambda,\theta}\left\{\theta\middle|\left(\min_{j:\lambda_j>0} y_{1j}\ldots \min_{j:\lambda_j>0} y_{sj}\right)\geq\theta y_k; X\lambda\leq x_k\lambda^T e\in\Gamma;\lambda\in\Re_+^n\right\}.$$

This model gives more efficient consistent estimators than models that impose less production assumptions, because it uses a larger subset of the production set.

3.6. Empirical selection procedure

A difficulty in implementing the transconvex model is the selection of the appropriate production assumptions and the specification of the corresponding transformations. Ideally, that issue is resolved on the basis of prior theoretical considerations or practical knowledge regarding the industry under evaluation. Section 3.5 illustrated that micro-economic theory can sometimes provide guidance for selecting the appropriate transformation. More specifically, that chapter demonstrated that particular exponential transformations can replace convexity for the production set with the less restrictive microeconomic assumption of convexity for the input sets and output sets.

However, if production knowledge is incomplete, the assumptions imposed may prove either erroneous or too general, potentially resulting in inconsistency or finite sample error respectively. If the production assumptions are erroneous, the estimators are statistically inconsistent, with a bias above true efficiency. That is because convex input-output vectors that in fact are outside the true production set are wrongly assumed attainable. By contrast, if the production assumptions are correct, the estimators are consistent. However, the estimators suffer from finite sample error. Moreover, the weaker the production assumptions (i.e. the stronger the transformations), the more error. Since erroneous production assumptions may result in inconsistency and assumptions that are too general in finite sample error, there is a need for determining the strongest yet correct set of production assumption. In the absence of complete prior production knowledge, we propose to use an empirical specification test to select the most progressive production assumptions supported by the data.

If a particular transformation convexifies the production set, TDEA gives statistically consistent estimates (provided free disposability (1.6) and the distribution assumptions (1.5), (1.8) and (1.15) are satisfied). Since FDH gives consistent estimates as well, the estimates of both models should converge as the sample size increases. However, if the transformation is too weak to convexify the production set, TDEA estimates are inconsistent with a positive bias and the estimates do not converge. In large enough samples, this insight can be used to test whether a particular transformation convexifies the production set.

For example, following Banker (1993), the following F-statistic can be used:

$$\textbf{(3. 61)}\quad F = \frac{\sum_{j=1}^{n}\left(1 - 1/\hat{\theta}^{G,H}_{TDEA,j}\right)^2 / n}{\sum_{j=1}^{n}\left(1 - 1/\hat{\theta}_{FDH,j}\right)^2 / n} \sim f_{n,n,1-\alpha}.$$

The rationale behind this test statistic is as follows. If the inefficiencies (i.c. $(1-\theta_j)$) are considered as identically and independently distributed random variables with a half-normal distribution, average squared inefficiency is a random variable with a chi-squared distribution. Hence, the average squared inefficiency *estimate* of a consistent model asymptotically follows a chi-squared distribution. In addition, the ratio of two independent chi-squared distributed random variables gives a random variable with a F-distribution. Hence, the ratio of the average squared inefficiency estimates of two consistent models asymptotically follows a F-distribution (see Banker (1993), p.

1272). As discussed above, if the transformations concavify the frontier, both FDH and TDEA are consistent. Therefore, the F-statistic can be used as an asymptotic test statistic for testing whether a particular transformation concavifies the frontier[47].

The maintained assumption of half-normality can be restrictive in some research environments. This assumption is a subject of considerable debate, and alternative test statistics have been proposed, including Kolmogorov-Smirnov tests (Banker (1993)) and t-tests (Kittelsen (1997)). Unfortunately, the finite sample performance of the competing tests is unclear. However, from the evidence available (e.g. Kittelsen, 1997), it appears that Banker's F-test performs relatively well under general conditions. In addition, the F-test demonstrably is asymptotically correct. Therefore, we decided for Banker's F-test in this study.

Significant deviations of TDEA estimates from the corresponding FDH estimates indicate that a stronger transformation is required to convexify the production set. Starting from the BCC model, production assumptions can be weakened by selecting more progressive domain and range transformations until the test statistic reaches the critical value corresponding to the desired level of significance.

For example, comparing BCC inefficiencies with the inefficiencies of the FDH model can test for convexity[48]. A difficulty in implementing the above approach is that the number of alternative transformations available is practically infinite. To circumvent this problem, we propose to focus on the following exponential transformations:

(3. 62) $G(Y) = e^{-pY}, H(X) = e^{pX} \qquad p > 0.$

The corresponding exponential TDEA model amounts to:

(3. 63) $$\hat{\theta}^{G,H}_{TDEA\,k} = \max_{\lambda,\theta}\left\{\theta \,\middle|\, \log\left(e^{-pY}\lambda\right)^{\frac{1}{p}} \geq \theta y_k;\ \log\left(\left(e^{pX}\lambda\right)^{\frac{1}{p}}\right) \leq x_k;\ \lambda^T e = 1;\ \lambda \in \Re^+_n\right\}.$$

The BCC and FDH models represent the limiting cases of this model for $p \to 0$ and $p \to \infty$ respectively.

Proof 3. 6 The limiting case for the generalized input (output) mean function for $p \to 0$ amounts to the convex input (output) combination (Hardy *et al.*, 1952):

$$\lim_{p\to 0}\left[\log\left(\left(e^{pX}\lambda\right)^{\frac{1}{p}}\right)\right] = X\lambda;$$

---

[47] The F-test was originally proposed for the purpose of comparing the efficiency of different and exclusive subsets of observations, keeping the model specification constant. By contrast, we propose the F-test for the purpose of testing the goodness of a particular model specification, keeping the data set constant. A complication in using the F-test for that purpose is that the evaluated models are both applied to the same data set, which introduces dependence for the average squared inefficiency estimates. To eliminate this problem, the average squared inefficiency estimates of both models can be computed over different (and exclusive) subsets of observations. For more details, see Kittelsen (1997).

[48] Convexity generally is not a property that is of direct interest to economists. However, as discussed above, convexity does exclude some interesting economic phenomena. For example, convexity excludes increasing returns-to-scale properties. For empirical tests that test returns-to-scale properties, see Kerstens and Vanden Eeckaut (1998).

$$\lim_{p\to 0}\left[\log\left((e^{-pY}\lambda)^{-\frac{1}{p}}\right)\right] = Y\lambda\,.$$

Consequently, the BCC model represents the limiting case for the exponential TDEA model for p→0. By analogy, the limiting case for p→∞ amounts to the maximum (minimum) inputs (outputs) of all referencing observations (Hardy *et al.*, 1952):

$$\lim_{p\to\infty}\left[\log\left((e^{pX}\lambda)^{\frac{1}{p}}\right)\right] = \left(\max_{j\in\{j:\lambda_j>0\}} x_{1j},\ldots,\max_{j\in\{j:\lambda_j>0\}} x_{nj}\right)^T;$$

$$\lim_{p\to\infty}\left[\log\left((e^{-pY}\lambda)^{-\frac{1}{p}}\right)\right] = \left(\min_{j\in\{j:\lambda_j>0\}} y_{1j},\ldots,\min_{j\in\{j:\lambda_j>0\}} y_{pj}\right)^T$$

Using this particular generalized mean function, TDEA selects as a reference unit the observation that achieves the highest radial output augmentation while consuming less or equal amounts of input than the evaluated unit, i.e. the FDH reference unit. □

In addition, the models that assume convex production sets and convex output sets discussed in Section 3.5 are special cases as well.

Focusing on the exponential domain transformation greatly simplifies the procedure for selecting the appropriate transformations, because it involves the single parameter *p* only. Moreover, the test statistic (3.63) is monotonically decreasing in *p*.

Proof 3. 7 For all $s \geq p$, the transformations $G'(y) = e^{-sy}$ and $H'(x) = e^{sx}$ are stronger than transformations $G(y) = e^{-py}$ and $H(x) = e^{px}$, because:

$$\begin{pmatrix}\log\left((e^{-sY}\lambda)^{-\frac{1}{s}}\right)\\ -\log\left((e^{sX}\lambda)^{\frac{1}{s}}\right)\end{pmatrix} \leq \begin{pmatrix}\log\left((e^{-pY}\lambda)^{-\frac{1}{p}}\right)\\ -\log\left((e^{pX}\lambda)^{\frac{1}{p}}\right)\end{pmatrix} \quad \forall (Y,X)\in T;\lambda^T e = 1;\lambda\geq 0;s\geq p\,.$$

Therefore, $\hat{T}^{G',H'}_{TDEA} \subseteq \hat{T}^{G,H}_{TDEA}$ and $\hat{\theta}^{G',H'}_{TDEA,k} \leq \hat{\theta}^{G,H}_{TDEA,k}$. Consequently, stronger transformations involve a lower test statistic than weaker ones:

$$\frac{\sum_{j=1}^{n}\left(1-1/\hat{\theta}^{G',H'}_{TDEA,j}\right)^2}{\sum_{j=1}^{n}\left(1-1/\hat{\theta}_{FDH,j}\right)^2} \leq \frac{\sum_{j=1}^{n}\left(1-1/\hat{\theta}^{G,H}_{TDEA,j}\right)^2}{\sum_{j=1}^{n}\left(1-1/\hat{\theta}_{FDH,j}\right)^2}.\square$$

Therefore, selecting the most progressive frontier supported by the data amounts to lowering *p* until the sum of squared inefficiencies exceeds the critical value.

Therefore, the exponential domain transformation induces a gradual transition between the FDH technology and the BCC technology, imposing increasingly progressive production assumptions as p decreases.

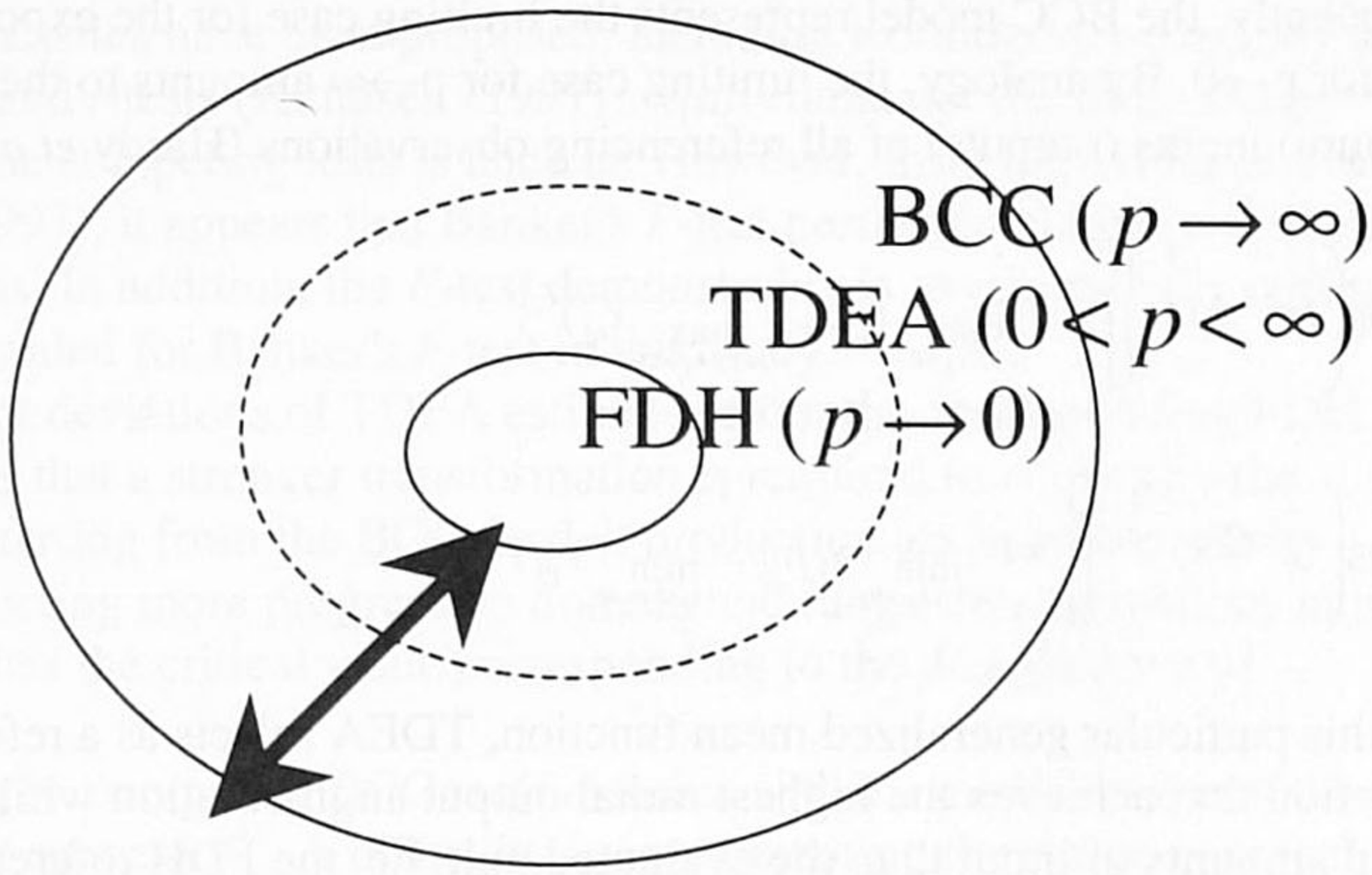


Figure 3. 15 FDH and BCC models as special cases of the exponential TDEA model

3.7. Simulating statistical properties

As demonstrated in Section 3.4, the transconvex model can consistently estimate non-convex production sets. However, statistical consistency relates to large sample performance only. Transforming input-output variables can increase small sample error and lower the rate of convergence. Unfortunately, no analytical solutions are available for the finite sample properties of DEA estimators. Consequently, we are not able to formally derive conditions under which TDEA improves efficiency estimation. To analyze the finite sample performance, we performed a series of Monte-Carlo simulations.

A difficulty of using simulated data is the selection of the appropriate data generating process (henceforth DGP) and the appropriate range for varying the parameters of the DGP. If the simulations do not represent a wide range of real-life research environments, the simulation results can have little significance for research practice. Unfortunately, to the best of our knowledge, there currently is no generally accepted framework for performing simulation studies in DEA. Therefore, we adopt a heuristic approach in this study. Specifically, we use a DPG that is relatively simple, and hence limits the computational burden, while allowing us to answer our research questions.

In the simulations, all units were assumed to face a common production frontier that maps two inputs (x_{1j},x_{2j}) into a single output (y_j). The following production frontier model represents that technology:

(3. 64) $y_j = f(x_{1j}, x_{2j})e^{-u_j}$;

$f(x_{1j}, x_{2j}) = 10x_{1j}^{\beta}x_{2j}^{\beta}$;

$u_j \sim |N(0, \sigma_u)|$.

$j = 1,\ldots,n$

The inputs were drawn from a bivariate lognormal distribution with (for both inputs) mean 100 and standard deviation 10 and coefficient of correlation 0.9. The log-normal distribution was selected to prevent negative values and moreover reflect the skewness of the size distribution typically encountered in empirical studies. Similarly, the correlation was specified to reflect the high degree of correlation between inputs typically encountered in empirical studies. Output was computed by first calculating the frontier outputs from the inputs and subsequently multiplying the frontier output with the efficiency percentage. The log-efficiencies (u_j) follow a half normal distribution, i.e. they are the absolute values of a variable with mean 0 and standard deviation σ_u. Under this assumption, log-efficiency has an average value of $(2/\pi)^{0.5}\sigma_u$ and a standard deviation of $((\pi\text{-}2)/\pi)^{0.5}\sigma_u$.

We investigated the influence of three parameters of DGP (3.64): the sample size (n), the standard deviation of the inefficiency term (σ_u), and the coefficient β. Other parameters could have been considered as well. However, we expect these parameters to be the main determinants of the goodness of the competing models. Recall that the goodness of the competing models involves a difficult trade-off between finite sample error and specification error. Finite sample error arises if the observations in the data set do not give a complete representation of the production possibilities. This is especially true is the sample is small (i.e. n is low) or if many observations are inefficient (i.e. σ_u is high). In addition, specification error arises if wrong production assumptions are imposed, allowing for unfeasible reference points. In the above DGP, the coefficient β determines whether or not the production set is convex. Convexity is satisfies if $\beta \leq 0.5$ and violations occur if $\beta > 0.5$.
Starting point for the simulations was a scenario with 200 observations, a 5 percent average inefficiency level and coefficient value β= 0.75. We subsequently varied systematically the values for n, σ_u, and β. More specifically, we considered samples of 50, 200 and 500 observations, average inefficiency levels of 2, 5 and 10 percent and values of 0.5, 0.75 and 1.0 for β.

We compared the BCC model (1.10) and the FDH model (3.2) with two variants of the transconvex model (3.22): a theory-based model and a data-based model. First, we analysed a model that uses a data transformation that is known a priori to convexify the production set. More specifically, we used the following data transformations:

(3. 65) $H(x_{1j}) = x_{1j}^{2\beta}$;

$H(x_{2j}) = x_{2j}^{2\beta}$;

$G(y_j) = y_j$.

These transformations can be demonstrated to convexify the production set used in the data generating process[49]. Second, we analyzed a model that selects the most progressive transformation supported by the data, using an empirical specification test. More particularly, the model uses Banker's F-test (3.61) to select the most progressive exponential domain transformation (3.62).

Each experiment consisted of generating a set of artificial data from the above data-generating process, and employing these data to create the inefficiency estimators from each of the three competing models. For each of the combinations of sample size, inefficiency level and production technology, 1000 experiments were done. Next, these estimates were used to gauge the sampling distribution properties of the competing estimators. We considered thee sample statistics: bias (BIAS), standard deviation (STD), and root mean squared error (RMSE). These statistics were computed as follows:

$$\textbf{(3. 66)}\quad BIAS = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j)\,;$$

$$\textbf{(3. 67)}\quad STD = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \bar{\hat{\theta}}_j)^2\,;$$

$$\textbf{(3. 68)}\quad RMSE = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j)^2\,.$$

Appendix 3.1 presents the GAUSS computer program used for the computations.

Figure 3.16 displays the sample distribution of the estimation errors $(\hat{\theta}_j - \theta_j)$ for the four models for the base case (n=200, $E(\varepsilon_j)$=0.05, and β=0.75). The BCC errors are centered well above zero with a bias of 0.078, and are widely spread out with a standard deviation of 0.045. Since β=0.75, the production set in the DGP (3.64) is non-convex. Consequently, the BCC model gives inconsistent estimates. It selects reference units that lie outside the production set and consequently confuses inefficiency with non-convexity. However, the frontier is monotonously increasing and the observations satisfy the distribution assumptions (1.5), (1.8) and (1.15) Therefore, the FDH does give consistent estimates. However, that model suffers greatly from finite sample error, as is reflected by a –0.044 bias and a standard deviation of 0.029. As discussed above, a low data density near the frontier, which results in a lack of efficient refercne units, causes this. The theory-based model ($TDEA_T$) is also consistent, because the data transformation (3.62) convexifies the production set. However, it suffers less from finite sample error than the FDH, because it considers generalized mean values of multiple observations rather than non-composite units only. Therefore, $TDEA_T$ depends less on the data density near the frontier, and consequently gives less error in finite samples. This is reflected in a bias of only –0.008 and a standard deviation of only 0.011. These statistics are

---

[49] These transformations convexify the production set, as $G(y_j)=\alpha H(x_{1j})^{0.5}H(x_{2j})^{0.5}$ is a concave function of the transformed variables.

significantly better than the corresponding FDH and BCC statistics[50]. The performance of the data-based model ($TDEA_D$) falls somewhat short of that of the theory-based model. This is because the former may select data transformations that are either too weak to convexify the production set (hence introducing specification error), or stronger than required to do so (hence introducing finite sample error). Still, with a -0.026 bias and a 0.021 standard deviation, $TDEA_D$ does significantly better than the FDH and BCC models.

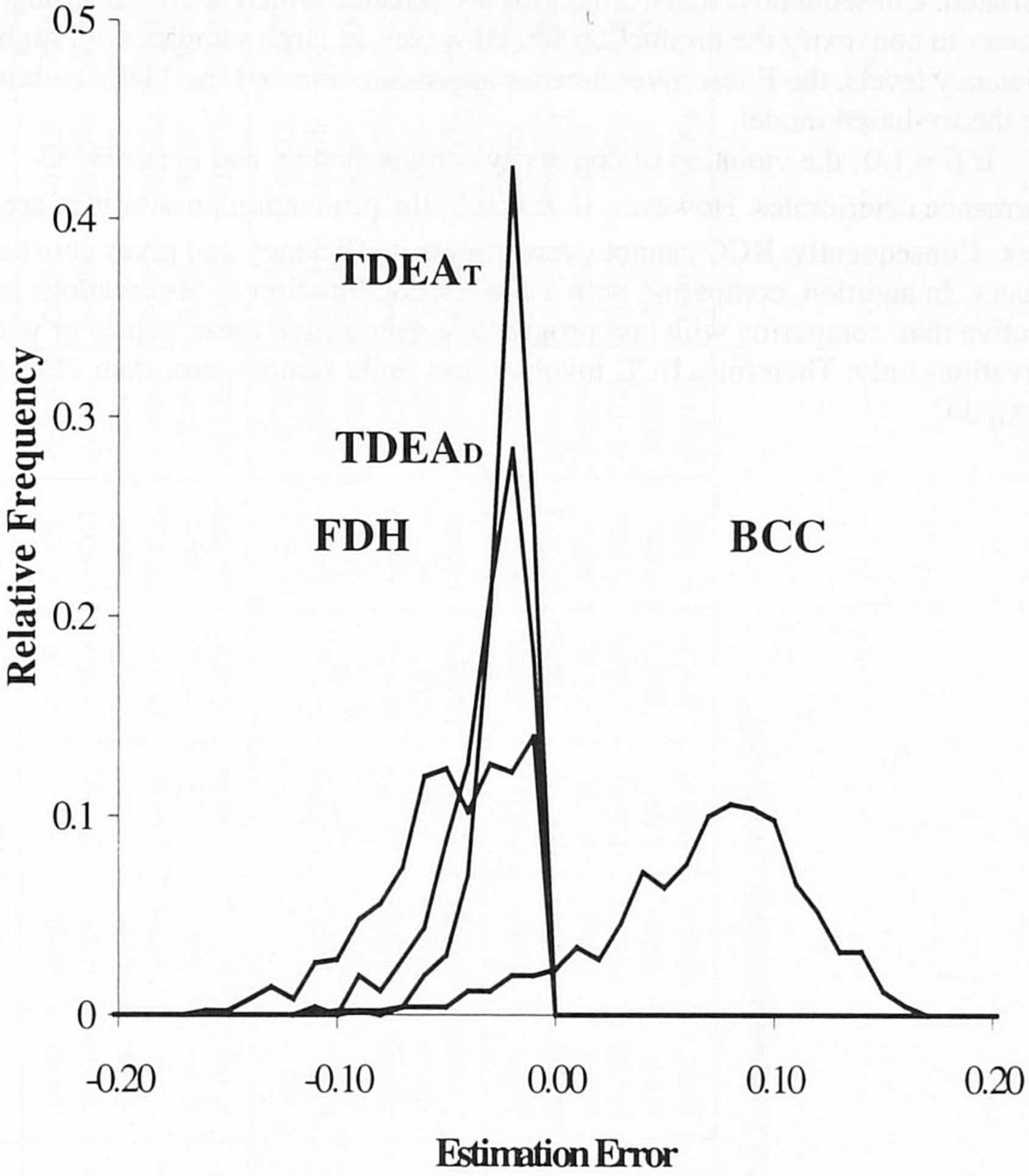


Figure 3. 16 Sample distribution of estimation error $(\theta_j - \theta_j)$ for the four competing models.

Varying sample size, inefficiency level and non-convexity leads to some changes in conclusions relative to the base case. Tables 3.2, 3.3 and 3.4 present the results for the

[50] We used a 1 percent significance level, which corresponds to critical values of 2.42, 1.71 and 1.42 for samples of 50, 200 and 500 observations respectively.

alternative settings considered. These tables display the average value and (in brackets) the standard deviations for the sample statistics computed over the 1000 replications. FDH and $TDEA_T$ performance improves as the sample size increases and as the inefficiency level decreases. If more units operate near the frontier, the finite sample error decreases. By contrast, BCC bias increases as the number of observations increases and inefficiency decreases, because the amount of convex combinations outside the production set increases as the number of efficient DMUs increases. The performance of $TDEA_D$ depends on the approximating power of the F-test. In small samples and with low inefficiency levels, the rejection rates of that test are inflated. Consequently, transformations are selected which are often stronger than necessary to convexify the production set. However, in large samples with high inefficiency levels, the F-test gives a better approximation and the $TDEA_D$ converges to the theory-based model.

If $\beta = 1.0$, the violation of convexity is more severe, and hence BCC performance deteriorates. However, if $\beta = 0.5$, the production possibilities are convex. Consequently, BCC cannot overestimate inefficiency and gives consistent estimates. In addition, comparing with a convex combination of observations is less restrictive than comparing with less progressive generalised mean values or with observations only. Therefore, BCC involves less finite sample error than FDH and $TDEA_D$ do.

		n=50				n=200				n=500			
		FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$
$E(\varepsilon_j)$ = 2%	BIAS	-0.021 (0.004)	-0.007 (0.002)	-0.020 (0.004)	-0.007 (0.002)	-0.019 (0.003)	-0.003 (0.001)	-0.016 (0.005)	-0.003 (0.001)	-0.016 (0.003)	-0.002 (0.001)	-0.012 (0.006)	-0.002 (0.001)
	STD	0.015 (0.003)	0.008 (0.004)	0.014 (0.003)	0.008 (0.004)	0.013 (0.003)	0.005 (0.003)	0.011 (0.004)	0.005 (0.003)	0.011 (0.002)	0.003 (0.003)	0.008 (0.004)	0.003 (0.003)
	RMSE	0.025 (0.004)	0.011 (0.004)	0.024 (0.004)	0.011 (0.004)	0.023 (0.003)	0.006 (0.003)	0.020 (0.006)	0.006 (0.003)	0.020 (0.003)	0.003 (0.003)	0.015 (0.006)	0.003 (0.003)
$E(\varepsilon_j)$ = 5%	BIAS	-0.045 (0.009)	-0.016 (0.007)	-0.036 (0.012)	-0.016 (0.007)	-0.039 (0.006)	-0.007 (0.003)	-0.020 (0.009)	-0.007 (0.003)	-0.032 (0.005)	-0.004 (0.005)	-0.008 (0.024)	-0.004 (0.005)
	STD	0.031 (0.005)	0.017 (0.006)	0.025 (0.007)	0.017 (0.006)	0.025 (0.005)	0.009 (0.006)	0.016 (0.008)	0.009 (0.006)	0.020 (0.004)	0.006 (0.005)	0.010 (0.022)	0.006 (0.005)
	RMSE	0.055 (0.009)	0.024 (0.007)	0.044 (0.012)	0.024 (0.007)	0.047 (0.007)	0.012 (0.006)	0.026 (0.011)	0.012 (0.006)	0.038 (0.005)	0.007 (0.005)	0.014 (0.024)	0.007 (0.005)
$E(\varepsilon_j)$ = 10%	BIAS	-0.082 (0.013)	-0.031 (0.010)	-0.051 (0.033)	-0.031 (0.010)	-0.062 (0.009)	-0.013 (0.005)	-0.017 (0.010)	-0.013 (0.005)	-0.048 (0.007)	-0.007 (0.004)	-0.008 (0.005)	-0.007 (0.004)
	STD	0.052 (0.019)	0.032 (0.015)	0.042 (0.044)	0.032 (0.015)	0.038 (0.009)	0.018 (0.012)	0.019 (0.012)	0.018 (0.012)	0.030 (0.006)	0.011 (0.010)	0.012 (0.010)	0.011 (0.010)
	RMSE	0.097 (0.015)	0.045 (0.017)	0.067 (0.054)	0.045 (0.017)	0.073 (0.011)	0.023 (0.012)	0.026 (0.015)	0.023 (0.012)	0.057 (0.008)	0.014 (0.010)	0.014 (0.011)	0.014 (0.010)

Table 3.2 Sample statistics for β=0.5

		n=50				n=200				n=500			
		FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$
$E(\varepsilon_j)$ = 2%	BIAS	-0.021 (0.003)	0.050 (0.024)	-0.020 (0.003)	-0.008 (0.002)	-0.020 (0.004)	0.096 (0.024)	-0.019 (0.023)	-0.004 (0.004)	-0.018 (0.004)	0.141 (0.029)	-0.010 (0.003)	-0.002 (0.001)
	STD	0.015 (0.003)	0.028 (0.012)	0.015 (0.003)	0.008 (0.003)	0.014 (0.003)	0.043 (0.017)	0.016 (0.023)	0.005 (0.004)	0.012 (0.003)	0.048 (0.012)	0.008 (0.003)	0.002 (0.001)
	RMSE	0.026 (0.004)	0.064 (0.026)	0.025 (0.004)	0.011 (0.003)	0.025 (0.004)	0.106 (0.024)	0.025 (0.023)	0.006 (0.004)	0.021 (0.004)	0.150 (0.029)	0.013 (0.004)	0.003 (0.001)
$E(\varepsilon_j)$ = 5%	BIAS	-0.048 (0.008)	0.025 (0.021)	-0.042 (0.014)	-0.017 (0.006)	-0.044 (0.008)	0.078 (0.027)	-0.026 (0.013)	-0.008 (0.003)	-0.037 (0.005)	0.117 (0.024)	-0.015 (0.018)	-0.004 (0.002)
	STD	0.035 (0.006)	0.040 (0.014)	0.033 (0.032)	0.018 (0.008)	0.029 (0.006)	0.045 (0.012)	0.021 (0.008)	0.011 (0.008)	0.024 (0.006)	0.048 (0.012)	0.020 (0.031)	0.007 (0.008)
	RMSE	0.059 (0.008)	0.052 (0.016)	0.055 (0.033)	0.025 (0.009)	0.053 (0.009)	0.091 (0.026)	0.035 (0.013)	0.014 (0.009)	0.045 (0.007)	0.127 (0.024)	0.029 (0.033)	0.008 (0.008)
$E(\varepsilon_j)$ = 10%	BIAS	-0.092 (0.015)	0.007 (0.029)	-0.066 (0.029)	-0.032 (0.013)	-0.074 (0.011)	0.058 (0.033)	-0.025 (0.016)	-0.016 (0.006)	-0.059 (0.008)	0.100 (0.025)	-0.002 (0.022)	-0.008 (0.003)
	STD	0.063 (0.013)	0.050 (0.017)	0.050 (0.023)	0.033 (0.018)	0.047 (0.010)	0.054 (0.016)	0.028 (0.013)	0.024 (0.014)	0.038 (0.010)	0.050 (0.012)	0.019 (0.012)	0.011 (0.007)
	RMSE	0.112 (0.017)	0.057 (0.020)	0.085 (0.029)	0.047 (0.021)	0.088 (0.013)	0.083 (0.028)	0.040 (0.016)	0.030 (0.015)	0.072 (0.012)	0.114 (0.021)	0.021 (0.018)	0.015 (0.008)

Table 3.3 Sample statistics for β=0.75

		n=50				n=200				n=500			
		FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$	FDH	BCC	$TDEA_D$	$TDEA_T$
$E(\varepsilon_j)$ = 2%	BIAS	-0.020 (0.004)	0.173 (0.081)	-0.020 (0.004)	-0.009 (0.002)	-0.020 (0.003)	0.314 (0.087)	-0.019 (0.004)	-0.004 (0.001)	-0.010 (0.002)	0.431 (0.088)	-0.010 (0.002)	-0.001 (0.001)
	STD	0.015 (0.003)	0.100 (0.030)	0.015 (0.003)	0.009 (0.003)	0.015 (0.002)	0.122 (0.028)	0.014 (0.003)	0.004 (0.003)	0.007 (0.001)	0.139 (0.036)	0.007 (0.001)	0.002 (0.001)
	RMSE	0.026 (0.004)	0.201 (0.083)	0.025 (0.004)	0.013 (0.004)	0.025 (0.003)	0.341 (0.087)	0.024 (0.004)	0.006 (0.003)	0.012 (0.002)	0.454 (0.088)	0.012 (0.002)	0.002 (0.001)
$E(\varepsilon_j)$ = 5%	BIAS	-0.051 (0.009)	0.143 (0.070)	-0.047 (0.013)	-0.018 (0.006)	-0.047 (0.007)	0.284 (0.085)	-0.036 (0.013)	-0.008 (0.003)	-0.040 (0.006)	0.416 (0.101)	-0.017 (0.025)	-0.005 (0.002)
	STD	0.037 (0.008)	0.098 (0.028)	0.035 (0.010)	0.018 (0.008)	0.032 (0.006)	0.127 (0.033)	0.027 (0.009)	0.011 (0.007)	0.027 (0.006)	0.139 (0.034)	0.021 (0.010)	0.008 (0.007)
	RMSE	0.063 (0.011)	0.175 (0.070)	0.059 (0.014)	0.025 (0.010)	0.056 (0.007)	0.318 (0.086)	0.045 (0.013)	0.014 (0.007)	0.048 (0.008)	0.440 (0.101)	0.034 (0.018)	0.009 (0.007)
$E(\varepsilon_j)$ = 10%	BIAS	-0.096 (0.016)	0.109 (0.078)	-0.072 (0.027)	-0.035 (0.012)	-0.083 (0.014)	0.257 (0.084)	-0.030 (0.037)	-0.016 (0.006)	-0.068 (0.011)	0.385 (0.104)	-0.004 (0.042)	-0.009 (0.005)
	STD	0.068 (0.014)	0.104 (0.032)	0.053 (0.017)	0.035 (0.016)	0.055 (0.012)	0.128 (0.033)	0.038 (0.019)	0.022 (0.017)	0.043 (0.010)	0.146 (0.039)	0.032 (0.017)	0.015 (0.014)
	RMSE	0.118 (0.019)	0.156 (0.073)	0.091 (0.027)	0.050 (0.019)	0.099 (0.016)	0.288 (0.082)	0.057 (0.027)	0.027 (0.017)	0.081 (0.012)	0.415 (0.100)	0.050 (0.025)	0.018 (0.014)

Table 3.4 Sample statistics for β=1.0

3.8. An empirical investigation into the production technology of bank branches

The empirical research into the productivity and efficiency of financial institutions has seen extensive application of DEA. For example, in a survey of frontier efficiency studies, Berger and Humphrey (1997) found that 69 out of 122 studies that compare firms use DEA, whereas 60 use parametric estimation techniques. In addition, 12 out of 13 studies focusing on branches use DEA, whereas only one study uses parametric techniques.

As discusssed in the previous above, DEA involves a difficult trade-off between the potential finite sample error associated with general DEA models, such as the FDH model, and the potential specification error associated with more specific models, such as the BCC and the CCR model. We think this trade-off is especially relevant for bank branch applications. Such applications typically study the branches of a single bank, and involve small or medium sized samples. Therefore, the FDH model can involve substantial finite sample error. In addition, there is good reason to assume that convexity is violated for the production technology for bank branches, because of increasing returns-to-scale for small-sized branches. Therefore, BCC and CCR may involve specification error.

This chapter presents the results of a real-life DEA study that includes an application of the transconvex model with convex input sets (CIS) discussed in Section 3.6. The study forms part of on-going research of the branch network of a large Dutch bank. The main activities of this bank's branch network are corporate and retail banking. Apart from traditional bank services, this bank actively offers insurance services in all relevant market segments. The network currently consists of 339 branches. In this study, we used management accounting information of 314 of those branches over 1997[51].

Variable selection

A number of DEA studies have analysed bank branch performance (Oral, *et al.* (1992); Al-Faraj, *et al.* (1993); Drake and Howcroft (1994); Sherman and Ladino (1995); Athanassopoulos (1997); and Schaffnit, *et al.* (1997)). Unfortunately, there is no generally accepted model for the technology of providing bank products and services. In general, the technology is represented by certain outputs (e.g. loans, mortgages and insurance), produced by certain inputs (e.g. personnel, workspace and computers) in a microenvironment, that is characterised by a certain number of environmental variables (e.g. number of competitors, consumer demographics, and the prices for the inputs and outputs). Figure 3.17 illustrates this general representation of the bank branch technology. Different studies focus on different output, input and environmental variables, depending on the research purpose, the banking theory the researchers adhere to, and the data available.

[51] The remaining 25 branches where not fully operational during 1997.

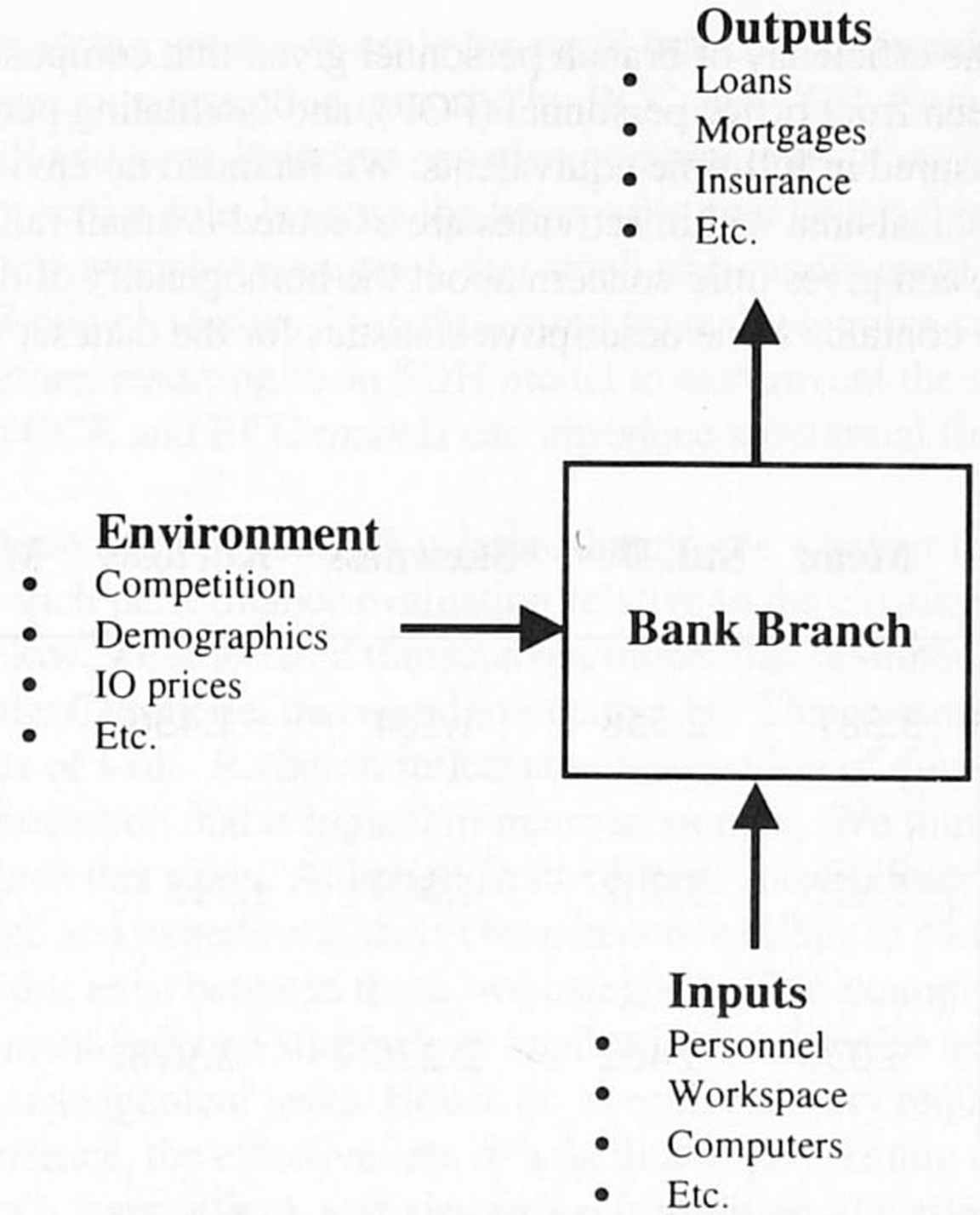


Figure 3. 17 A general representaion of the bank branch technology, involving inputs, outputs and environmental variables

The purpose of this study was to develop a framework for evaluating the revenue generating efficiency of bank branch personnel. For that purpose, we used a single output, total revenues (TR), which includes capital revenues and commissions on insurance, brokerage and payment services, all of which are measured in millions of US dollars (mUSD). Total revenues uses output prices to aggregate multiple outputs into a single revenue measure, which helps to reduce the dimensionality and the associated finite sample error. Note that we excluded extraordinary revenues so as to increase the robustness of DEA with respect to data variations[52]. Since we study branches of the same bank, we assume that the variations in the physical facilities (workspace, computers, etc.) are minimal, and we focus on personnel input only. This approach was also taken by Schaffnit *et al.* (1997). Revenues are mainly generated by front-office personnel, e.g. account managers and specialist advisors. However, facilitating personnel, including general management, administrators and other facilitating personnel, supports front-office operations by providing management and administrative support and maintaining the physical facilities. It is important to distinguish between these two types of personnel, so as to separate the effects of the composition of branch

[52] DEA results can be sensitive to data variations. Outlier detection procedure, robustness tests and stochastic DEA models can help to increase robustness (for more details, see Chapter 4). Another approach, adopted in this application, is to select economic variables that are minimally perturbed by incidental factors.

personnel from the efficiency of branch personnel given that composition. Therefore, we distinguish between front office personnel (FOP), and facilitating personnel (FAC), all of which are measured in full-time equivalents. We included no environmental variables since the geographical area within activities are executed is small (all branches operate in the Netherlands), and gives little concern about the homogeneity of the observations.

Table 3.5 contains some descriptive statistics for the data set[53].

	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
FOP (FTE)	3.387	2.558	1.184	1.436	0.000	13.000
FAC (FTE)	5.925	3.702	1.319	1.715	0.000	21.400
TR (mUSD)	3.038	3.462	2.275	3.918	0.379	43.036

Table 3.5 Descriptive statistics for the bank branch data set

Model selection

In addition to using different output, input and environmental variables, different studies use different assumptions with respect to the relationships between the variables (e.g. free disposability (FDH), convexity (BCC), and linear homogeneity (CCR)). As discussed above, imposing general assumptions only may result in finite sample error, whereas imposing erroneous assumptions can result in specification error. We think the resulting trade-off is especially relevant in bank branch studies.

First, there is reason to believe that convexity is violated, because of the limited opportunities for the division of labour at small-sized branches. Branch operations involve many different and complex tasks, including specialist account management, general management activities, and administration. The diversity and complexity of these tasks calls for specialised personnel with specialised knowledge and experience. We have already discussed the distinction between front-officers and facilitators. In addition, front office personnel can further specialise in different market segments or advice areas, and facilitating personnel can specialise in administrative tasks or general management tasks. Unfortunately, dividing labour requires a minimal number of employees. Consequently, small-sized branches typically employ generalists that can be employed for a broad range of tasks, but typically lack specialized knowledge and experience. This reduces the revenue potential of small-sized branches relative to large branches. Ideally, the heterogeneity of personnel would be reflected in the variable selection. Unfortunately, the data set allows us to separate between front-officers and facilitators only, and we have to assume homogeneity within these categories. Therefore,

[53] We subjected the original data to a linear transformation, because of the proprietary nature of the data. Altough the *absolute* magnitudes of the variables are lost, the transformation preserves the *relative* magnitudes.

we expect to find increasing returns-to-scale for small branches. Increasing returns-to-scale violate the convexity assumption imposed by BCC and CCR. Convex combinations of small and large branches can give combinations of revenues and personnel that are not achievable, because the large units may have achieved their revenues by employing specialist personnel, that small unit cannot employ.

Second, bank branch studies (including ours) typically involve small or medium-sized samples. Therefore, resorting to an FDH model to circumvent the specification error associated with CCR and BCC models can introduce substantial finite sample error.

Because of these complications, it is interesting to see whether the CIS model can improve bank branch performance evaluation relative to the existing models. For reasons discussed below, we selected a transconvex model that assumes that the input sets are convex, i.e. the CIS model discussed in Section 3.6. This assumption does not restrict the economies of scale. Rather, it reflects the assumption of diminishing marginal rates of technical substitution that is typical in microeconomics. We think that assumption is tenable in this study. Although front-officers and facilitators have specialised knowledge and experience, their competence overlaps to some extent, and substitution possibilities exist between these two categories. For example, administrators can replace account managers in a simple loan application, and senior account managers can perform general management tasks. However, as each category requires specialised knowledge and experience, the effectiveness of a facilitator performing a front office job is smaller than that of a front officer, and vice versa. In addition, the effectiveness becomes smaller if more personnel are substituted and facilitators (front-officers) are assigned specialist front office (support) tasks. For example, an administrator can handle a loan application, but he typically lacks the training and experience to give more specialised advice. Therefore, as the reduction of facilitators (front-officers) increases, the additional front-officers (facilitators) required to sustain revenues increases progressively. Consequently, the assumption of convexity seems tenable for the input sets, and the CIS could potentially improve performance evaluation.

Apart from theoretical and practical considerations, we analyzed the data to gauge the validity of the assumptions imposed by the alternative DEA models. More specifically, we estimated the following econometric frontier:

$$\textbf{(3. 69)}\quad TR_j = \alpha + \beta_1 FOP_j + \beta_2 FAC_j + \gamma_1 FOP_j^2 + \gamma_2 FAC_j^2 + \delta FOP_j \cdot FAC_j - u_j$$
$$u_j \sim \left| N(0,\sigma_u) \right| .$$

This model has a quadratic structure that is flexible enough to give a second-order approximation to an arbitrary twice-differentiable frontier[54]. In addition, it includes an inefficiency term with a one-sided, half-normal distribution. Unlike Stochastic Frontier Analysis models, (3.69) is deterministic by nature, and does not include a stochastic disturbance term, so as to make it better comparable with the DEA models, which do also have a deterministic nature[55].

---

[54] The translog specification is an example of the flexible functional forms that have been developed in econometric frontier estimation for relaxing the requirement for a complete specification of the functional form of the frontier (see Diewert and Wales (1987), and Pulley and Braunstein (1992), among others).

[55] By excluding noise, this specification follows the deterministic frontier tradition developed by Aigner and Chu (1968), Førsund and Jansen (1977) and Førsund and Hjalmarsson (1979a, 1979b),

The validity of the assumptions imposed by alternative DEA models can be gauged by testing restrictions on the regression coefficients in the above model[56]. These properties can be derived from the first and second order derivatives of the revenue function. More specifically, FDH imposes free disposability, which requires the first-order derivatives

$$\textbf{(3. 70)}\quad \left(\frac{\partial TR_j}{\partial FOP_j}\right)=\beta_1+2\gamma_1 FOP_j+\delta FAC_j \quad \text{and}$$

$$\textbf{(3. 71)}\quad \left(\frac{\partial TR_j}{\partial FAC_j}\right)=\beta_2+2\gamma_2 FAC_j+\delta FOP_j$$

to be non-negative for all $j \in D$. CIS imposes convexity for the input sets in addition to free disposability, which requires the second-order bordered leading principal minor (i.e. the determinant of the bordered Hessian matrix)

$$\textbf{(3. 72)}\quad 2\left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right)\left(\frac{\partial TR_j}{\partial FOP_j}\right)\left(\frac{\partial TR_j}{\partial FAC_j}\right)-\left(\frac{\partial^2 TR_j}{\partial FOP_j^2}\right)\left(\frac{\partial TR_j}{\partial FAC_j}\right)^2-\left(\frac{\partial^2 TR_j}{\partial FAC_j^2}\right)\left(\frac{\partial TR_j}{\partial FOP_j}\right)^2 =$$

$$2\delta(\beta_1+2\gamma_1 FOP_j+\delta FAC_j)(\beta_2+2\gamma_2 FAC_j+\delta FOP_j)-2\gamma_1(\beta_2+2\gamma_2 FAC_j+\delta FOP_j)^2-2\gamma_2(\beta_1+2\gamma_1 FOP_j+\delta FAC_j)^2$$

to be positive for all $j \in D$[57]. Finally, BCC imposes convexity, which requires the Hessian matrix to be negative semi-definite, i.e. the second-order partial derivatives

$$\textbf{(3. 73)}\quad \left(\frac{\partial^2 TR_j}{\partial FOP_j^2}\right)=2\gamma_1 \quad \text{and}$$

$$\textbf{(3. 74)}\quad \left(\frac{\partial^2 TR_j}{\partial FAC_j^2}\right)=2\gamma_2$$

are assumed non-positive and the Hessian determinant

---

among others, rather than the stochastic frontier tradition developed by Aigner, Lovell and Schmidt (1977), Meeusen and Van den Broeck (1977) and Battese and Corra (1977), among others.

[56] Obviously, the conclusions obtained will be condtional upon the maintained assumption that the econometric frontier (3.69) is flexible enough to give a global approximation of the true efficient frontier, and in addition that the inefficiencies have a half-normal distribution.

[57] The bordered Hessian matrix is given by:

$$\begin{bmatrix} \left(\frac{\partial^2 TR_j}{\partial FOP_j^2}\right) & \left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right) & \left(\frac{\partial TR_j}{\partial FOP_j}\right) \\ \left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right) & \left(\frac{\partial^2 TR_j}{\partial FAC_j^2}\right) & \left(\frac{\partial TR_j}{\partial FAC_j}\right) \\ \left(\frac{\partial TR_j}{\partial FOP_j}\right) & \left(\frac{\partial TR_j}{\partial FAC_j}\right) & 0 \end{bmatrix}.$$

$$\textbf{(3. 75)}\quad \left(\frac{\partial^2 TR_j}{\partial FOP_j^2}\right)\left(\frac{\partial^2 TR_j}{\partial FAC_j^2}\right)-\left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right)^2 = 4\gamma_1\gamma_2 - \delta^2$$

is assumed non-negative, for all $j \in D$[58]. For a more detailed discussion of necessary and sufficient conditions for convexity of the production set and its input sets, see Madden (1986).

Model (3.69) was estimated by maximum likelihood techniques. Table 3.6 displays the estimated coefficients and, in brackets, the associated p-values.

	Estimate (p-value)
α	3.977 (0.000)
β_1	0.947 (0.000)
β_2	0.010 (0.395)
γ_1	0.048 (0.012)
γ_2	0.017 (0.417)
δ	0.083 (0.000)

Table 3.6 Estimation results for the econometric frontier

These outcomes support our prior view on the branch technology. Revenues seem to be generated predominantly by front office personnel and not by facilitating personnel directly, as β_1 and γ_1 are positive and highly significant, whereas β_2 and γ_2 are insignificant. However, the facilitators do seem to substantially increase the revenue generating potential of front office employees, as δ is positive and highly significant. In addition, the data suggest increasing return-to-scale for front-office personnel, as γ_1 is significantly positive. As discussed above, these return-to-scale presumably are attributable to the limited opportunities for dividing labor at small branches.

The findings suggest that free disposability and convexity of input sets hold, whereas convexity is violated by the data. More specifically, the first-order derivatives and the second-order bordered leading principal minor are strictly positive for all observations. Imposing negative semi-definiteness on the Hessian matrix significantly

---

[58] The Hessian matrix is given by:

$$\begin{bmatrix} \left(\frac{\partial^2 TR_j}{\partial FOP_j^2}\right) & \left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right) \\ \left(\frac{\partial^2 TR_j}{\partial FOP_j \partial FAC_j}\right) & \left(\frac{\partial^2 TR}{\partial FAC_j^2}\right) \end{bmatrix}.$$

reduced the likelihood. This suggests that the FDH and the CIS assumptions hold, whereas, the BCC and CCR assumptions are violated.

Figures 3.18 and 3.19 illustrate these results Figureically. Figure 3.18 displays the estimated surface in three dimensions. The efficient surface clearly is monotonously increasing, but it is not concave, because of its positive curvature for small branches.

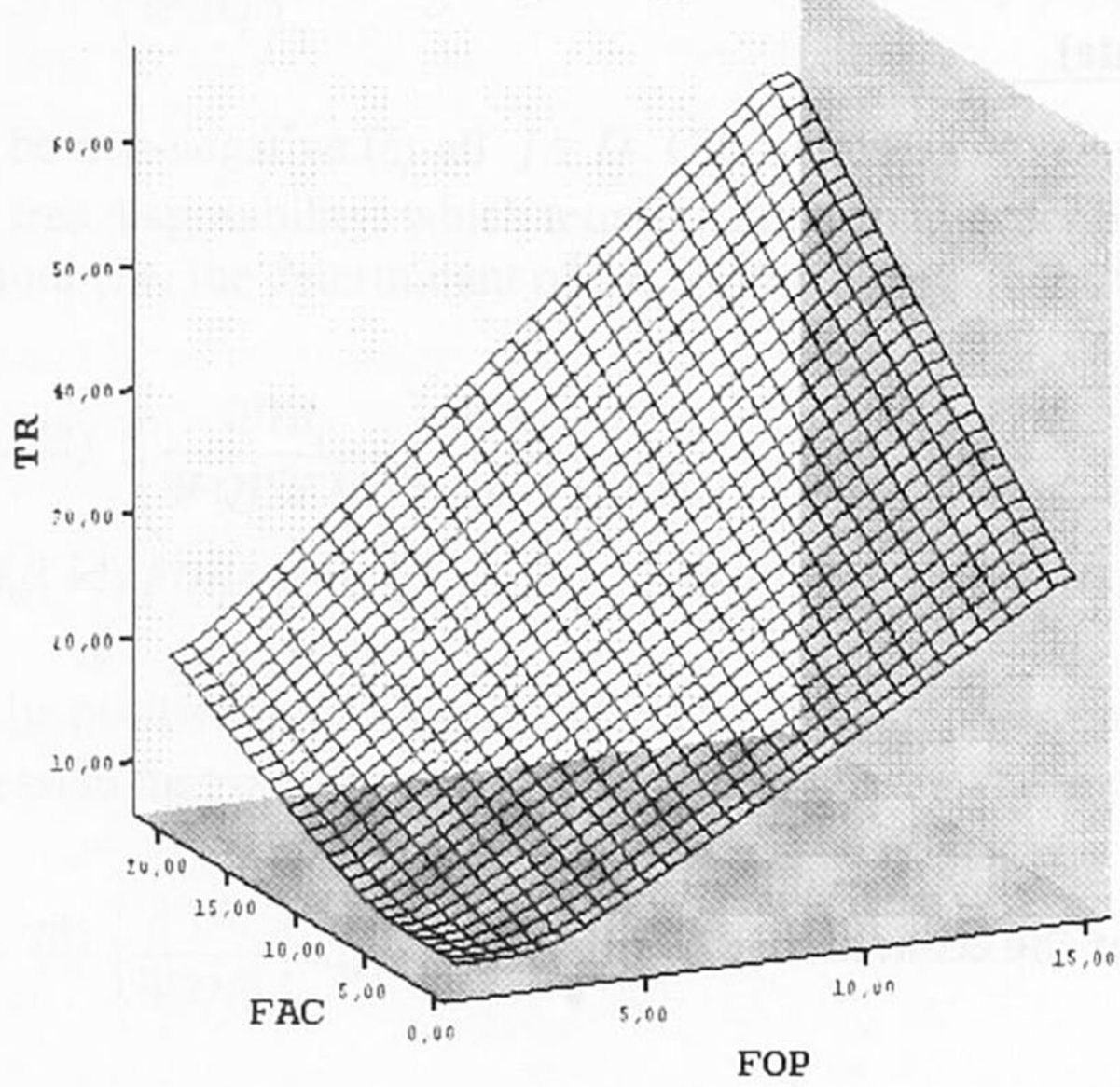


Figure 3. 18 A 3-dimensional Figure of the revenue frontier. The curvature at small personnel levels indicates substantial increasing returns-to-scale for small bank branches

Figure 3.19 displays some isoquants for the frontier. The isoquants are clearly convex. In addition, the isoquants are progressively closer to each other as the revenue level increases, which reflects the increasing returns-to-scale properties.

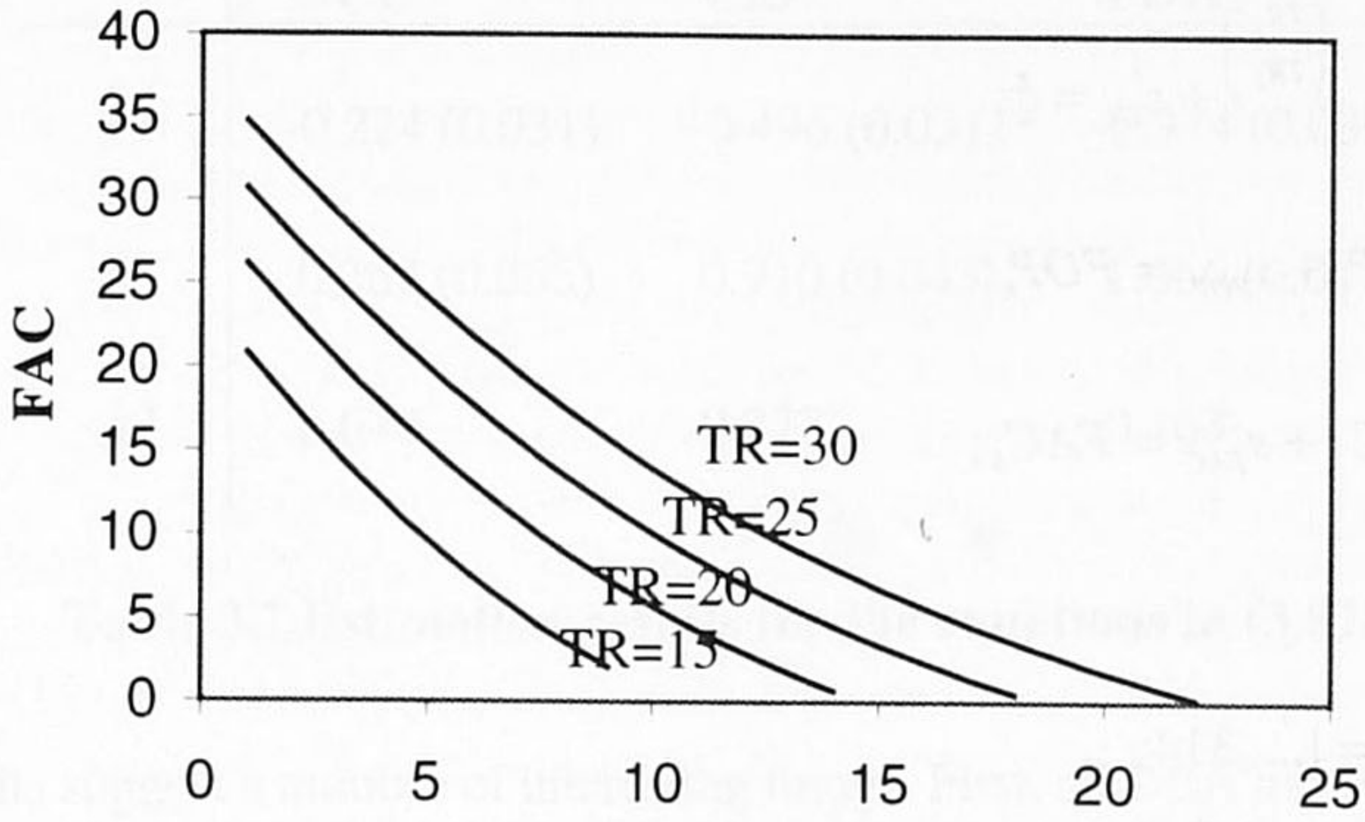


Figure 3. 19 Four isoquants of the revenue frontier. The isoquants are convex, which suggests that the TDEA model is correctly specified.

Still, these findings do not necessarily imply that TDEA is the best model, because, as discussed above, imposing additional assumptions can reduce the finite sample error associated with models that rely on minimal assumptions. To assess the goodness of the alternative DEA models, we compared the coefficients of correlation between the efficiency estimates of the econometric model and three alternative DEA models: FDH (3.2), CIS (3.47) and BCC (1.10).

Specifically, the following Mixed Integer Linear Programming gives the FDH estimate:

$$\textbf{(3. 76)} \quad \max_{\theta,\lambda,s'_{TR},s_{FOP},s_{FAC}} \quad \theta$$

$$\text{s.t.} \sum_{j=1}^{314} \lambda_j TR_j - s'_{TR} = \theta TR_k;$$

$$\sum_{j=1}^{314} \lambda_j FOP_j + s_{FOP} = FOP_k;$$

$$\sum_{j=1}^{314} \lambda_j FAC_j + s_{FAC} = FAC_k;$$

$$\sum_{j=1}^{314} \lambda_j = 1;$$

$$\lambda_j \in \{0,1\}\ j = 1,\ldots,314;$$

$$s'_{TR}, s_{FOP}, s_{FAC} \geq 0.$$

In addition, the following Linear Programming model was used to derive the CIS estimate:

(3. 77) $$\max_{\theta,\lambda,s'_{TR},s_{FOP},s_{FAC}} \xi$$

$$\text{s.t.}\sum_{\text{j=1}}^{314}\lambda_j e^{-p\left(\frac{TR_j}{TR_k}\right)} + s'_{TR} = \xi;$$

$$\sum_{j=1}^{314}\lambda_j FOP_j + s_{FOP} = FOP_k;$$

$$\sum_{j=1}^{314}\lambda_j FAC_j + s_{FAC} = FAC_k;$$

$$\sum_{j=1}^{314}\lambda_j = 1;$$

$$\lambda_j \geq 0 \;\; j = 1,\ldots,314;$$

$$s'_{TR}, s_{FOP}, s_{FAC} \geq 0.$$

The optimal solution is a transform of the CIS estimate. The estimate can be obtained as $\hat{\theta}_{CIS,k} = -\frac{1}{p}\ln\xi^*$. Finally, the follwing Linear Programming model gives theBCC estimates:

(3. 78) $$\max_{\theta,\lambda,s'_{TR},s_{FOP},s_{FAC}} \theta$$

$$\text{s.t.}\sum_{\text{j=1}}^{314}\lambda_j TR_j - s'_{TR} = \theta TR_k;$$

$$\sum_{j=1}^{314}\lambda_j FOP_j + s_{FOP} = FOP_k;$$

$$\sum_{j=1}^{314}\lambda_j FAC_j + s_{FAC} = FAC_k;$$

$$\sum_{j=1}^{314}\lambda_j = 1;$$

$$\lambda_j \geq 0 \;\; j = 1,\ldots,314;$$

$$s'_{TR}, s_{FOP}, s_{FAC} \geq 0.$$

To make the models better comparable, we transformed the parametric inefficiency measure u_j in a percentage efficiency measure using $\hat{\theta}_{EFE} = (TR_j + \hat{u}_j)/TR_j$.
Subsequently, we analyzed the comparative performance of the models by estimating the following regression equations:

(3. 79) $\hat{\theta}_{PFA,j} = \alpha_i + \beta_i\hat{\theta}_{i,j} + \varepsilon_j \quad j \in D; i \in \{\text{BCC,CIS,FDH}\}.$

These equations were estimated using the Ordinary Least Squares method. The CIS model is supported by the data if it has a higher coefficient of correlation than both the BCC and the FDH models. Table 3.7 displays the estimation output.

	BCC	CIS	FDH
α_i	-0.224 (0.031)	-0.496 (0.031)	-0.974 (0.036)
β_i	0.362 (0.055)	0.910 (0.043)	1.066 (0.047)
R^2	0.070	0.277	0.163

Table 3.7 Estimation results for the equations in (3.81)

The results suggest a number of interesting things. First, all DEA models give results that differ substantially from the results of the parametric model, with coefficients of correlation ranging from 0.070 to 0.273. These differences may be attributable to the fact that the quadratic function and the half-normal distribution are too restrictive to give a good global approximation, or alternatively to the fact that DEA models can suffer from substantial error in finite samples, because they impose minimal structure. Second, the CIS model gives the highest correlation (0.277), whereas the BCC gives the lowest correlation (0.07). These results support the prior view that the BCC is mis-specified because it relies on the erroneous assumption of convexity for the production set, and in addition that the FDH suffers seriously from small sample bias, because it imposes minimal assumptions, and hence requires more observations than available in this study. The CIS model has the highest correlation, because it does not impose the erroneous assumption of convexity, but does impose the correct assumption of convex input sets.

3.9. Concluding remarks

This chapter has demonstrated that the DEA methodology comprises a wide variety of alternative models that impose alternative sets of production assumptions. In addition, emprical specification tests have been developed that allow for the selection of the model that fits the data best. Finally, nonobserved DMUs can be added to the data set to include additional, local production information. These advances greatly extend the flexibility of the DEA methodology to deal with a variety of different problems.

Still, most DEA models maintain the assumption that the production possiblities are convex. Unfortunately, production characteristics such as increasing returns-to-scale, increasing marginal rates of substitution for the inputs, and increasing marginal rates of transformation for the outputs can violate this assumption. The FDH model can deal with such non-convexities, provided the production set satisfies free disposability. However, given its finite sample performance, it is desirable to develop more efficient non-convex models.

This chapter demonstrated how transforming the input-output variables can account for non-convex production technologies. TDEA gives consistent estimates for a broader set of production technologies than BCC and involves less finite sample error than the FDH. The Monte-Carlo simulations suggest that the reduction in estimation error relative to both BCC and FDH can be substantial.

Sometimes the appropriate transformations can be selected using prior theoretical considerations or practical knowledge of the industry under evaluation. For

example, the transconvex CIS, COS and CIOS models can impose the microeconomic assumption that the input and output sets are convex, as proposed by Petersen (1990). The transconvex models have some advantages relative to the Petersen model. First, they are demonstrably consistent with the assumptions of convexity for the input and/or output sets. Second, the transconvex models allow for the separate introduction and testing of convexity for the input sets and convexity for the output sets. Third, there is no need for the prior specification of the returns-to-scale properties. However, if return-to-scale information is available, it can be incorporated in the transconvex models. Finally, the transconvex models give feasible solutions, whereas the Petersen model may not have a feasible solution for non-increasing returns-to-scale technologies.

In addition, we proposed an empirical selection procedure for selecting the appropriate transformation if the convexifying transformations can not be determined a priori. That procedure allows for a gradual transition between BCC and FDH, models which represent the limiting cases of the exponential TDEA model. The simulations suggest that the empirical selection procedure can suffer from the poor finite sample performance of the embedded specification test. Still, it can substantially reduce bias and error variance relative to the classical models, especially in large samples and if convexity is severely violated.

The empirical application for bank branches demonstrates that TDEA is not just a theoretical possibility. Theoretical and practical arguments as well as empirical evidence suggest that the bank branch technology is characterized by substantial increasing returns-to-scale. Presumably, this phenomenon can be explained by the limited opportunities small branches have for employing specialist personnel. This finding conflicts with the assumption of convexity imposed by the BCC model, and implies that BCC gives inconsistent estimates. In addition bank branch studies typically involve small or medium sized samples. Therefore, the FDH model can involve substantial finite sample error. However, practical considerations and the data do support convexity of input sets, as is theoretically justified by the law of diminishing marginal rates of technical substitution. Therefore, in contrast to BCC, the CIS model does give consistent estimators. In addition, it suffers less from finite sample error than the FDH. Therefore, we conclude that the transconvex model is a potentially valuable addition to the research instruments currently available for evaluating bank branch performance.

The choice of the specification test used may prove critical to databased TDEA. However, relatively little is known about the finite sample performance of the existing tests. Moreover, the existing evidence suggests that these tests are sensitive to finite sample bias and dependence. Therefore, in our opinion, efforts should be directed at gathering systematic evidence on the finite sample performance of existing specification tests and bias correction methods, and at designing new ones.

A difficulty in simulating the goodness of the competing models is the absence of a generally accepted framework for performing simulation studies in DEA. Future research should focus on formulating data generating processes that give a good representation of real-life research environments.

Appendix 3.1

This appendix presents the GAUSS code for computing the models used in the simulations in section 3.7. The global variables correspond to the variables used in data generating process (3.66). Table 3.8 explains this correspondence.

GAUSS Global Variable	Symbol	Description
SampleSize	n	Sample size
Beta	β	Beta coefficient of the production frontier
StdIneff	σ_u	Standard deviation of inefficiencies

Table 3.8 Global variables used in the GAUSS code

GAUSS code

```
proc (3) = TDEA(SampleSize,Beta,StdIneff);

local RandomFile,Estimators,HelpFile,A,B,C,L,U,Lambda;
local Inefficiency,X,Y, Xtransformed;Bigstep,Smallstep;
local BIAS,STD,RMSE,X1,X2,F,P,I,T;

/********** BEGIN initialise local variables **************/

Estimators=zeros(SampleSize,4);
BIAS=zeros(1000,4);
STD=zeros(1000,4);
RMSE=zeros(1000,4);
P=zeros(1000,1);
F=zeros(1000,1);
U=1E256;
L=0;
__output=0;
_lpmin=0;
Bigstep=0.1;
Smallstep=0.01;

/*************** END initialise local variables *************/

T=1;
do while T .le 1000;
```

```
/*************** BEGIN simulation data set ***************/

   X1=(100+rndn(SampleSize,1)*10);
   X2=(100+0.9*(X1-100)+rndn(SampleSize,1)*sqrt(19);
   X=exp(X1~X2);
   Inefficiency=1-exp(-abs(rndn(SampleSize,1)* StdIneff));
   Y=10.*(X1.^Beta).*(X2.^Beta).*(1-Inefficiency);
   RandomFile=q~X1~X2;

/*************** END simulation data set ******************/

/*************** BEGIN FDH-model ********************/

   I=1;
   do while I .le SampleSize;

         HelpFile=selif(RandomFile,(RandomFile[.,2 3] .lt RandomFile[I,2 3]);
         Estimators[I,1]=(maxc(HelpFile[.,1])-
         RandomFile[I,1])./RandomFile[I,1];
         I=I+1;

   endo;
   BIAS[T,1]=(sumc((Estimators[.,1]-Inefficiency))/SampleSize);
   STD[T,1]=stdc(Estimators[.,1]-Inefficiency);
   RMSE[T,1]=sqrt(BIAS[T,1]^2+STD[T,1]^2);

/******************* END FDH model **********************/

/****************** BEGIN BCC model *********************/

   I=1;
   do while I .le SampleSize;

         Library simplex;
         _lpcnst=ones(1,cols(X))~2~3;
         A = ( X~Y~ones(rows(X),1) )';
         B = (X[I,.]~Y[I]~1)';
         C = Y;
         {Lambda, Inefficiency, ReturnCode} = simplex(A, B, C,L, U);
         Estimators[I,2]=(Inefficiency-RandomFile[I,1])./RandomFile[I,1];
         I=I+1;

endo;
BIAS[T,2]=(sumc((Estimators[.,2]-Inefficiency))/SampleSize);
STD[T,2]=stdc(Estimators[.,2]-Inefficiency);
RMSE[T,2]=sqrt(BIAS[T,2]^2+STD[T,2]^2);
if sumc(Estimators[.,1].^2) .eq 0;

         F[T]=999;
```

```
    else;
            F[T]=sumc(Estimators[.,2].^2)/(sumc(Estimators[.,1].^2));
    endif;
/******************** END BCC model ************************/
/****************** BEGIN TDEA-D model ******************/
     if cdffc(F[T],SampleSize, SampleSize ) .ge 0.01;
            BIAS[T,3]=BIAS[T,2];
            STD[T,3]=STD[T,2];
            RMSE[T,3]=RMSE[T,2];
else;
            P[T]=0;
            do while cdffc(F[T],SampleSize,sampleSize) .lt 0.01;
                  P[T]=P[T]+Bigstep;
                  I=1;
                  do while I .le SampleSize;
                  library simplex;
                  _lpcnst=ones(1,cols(X))~2~3;
                  Xtransformed=exp(P[T]*X);
                  A = (Xtransformed~Y~ones(rows(X),1) )';
                  B = (Xtransformed[I,.]~Y[I]~1)';
                  C =Y;
                  {Lambda ,Inefficiency, ReturnCode} =
                  simplex(A, B, C, L, U);
                  Estimators[I,3]=(Inefficiency- Y[I])./Y[I];
                  I=I+1;
            endo;
            F[T]=sumc(Estimators[.,3].^2)/
            (sumc(Estimators[.,1].^2));
            endo;
            do while cdffc(F[T],n,n) .Gt 0.01;
                  P[T]=P[T]-Smallstep;
                  I=1;
                  do while I .le Sample Size;
                        library simplex;
                        _lpcnst=ones(1,cols(X))~2~3;
```

```
                    Xtransformed=exp(P[T]*X);
                    A =(Xtransformed~Y~ones(rows(X),1) )';
                    B =(Xtransformed[I,.]~Y[I]~1)';
                    C =Y;
                    {Lambda ,Inefficiency, ReturnCode} =
                    simplex(A, B, C, L, U);
                    Estimators[I,3]=(Inefficiency-Y[I])./Y[I];
                    I=I+1;

        endo;
        F[T]=sumc(Estimators[.,3].^2)/(sumc(Estimators[.,1].^2));
                    endo;
                    BIAS[T,3]=(sumc((Estimators[.,3]Inefficiency))/SampleSize);
                    STD[T,3]=stdc(Estimators[.,3]-Inefficiency);
                    RMSE[T,3]=sqrt(BIAS[T,3]^2+STD[T,3]^2);

    endif;

/*************** END TDEA-D model ******************************/

/*************** BEGIN TDEA-T model ****************************/

   I=1;
   do while I .le SampleSize;

        library simplex;
        _lpcnst=ones(1,cols(X))~2~3;
        Xtransformed=X.^(2*Beta);
        A =(Xtransformed~Y~ones(rows(X),1) )';
        B =(Xtransformed[I,.]~Y[I]~1)';
        C =Y;
        {Lambda ,Inefficiency, ReturnCode} = simplex(A, B, C, L, U);
        Estimators[I,4]=(Inefficiency-Y[I])./Y[I];
        I=I+1;

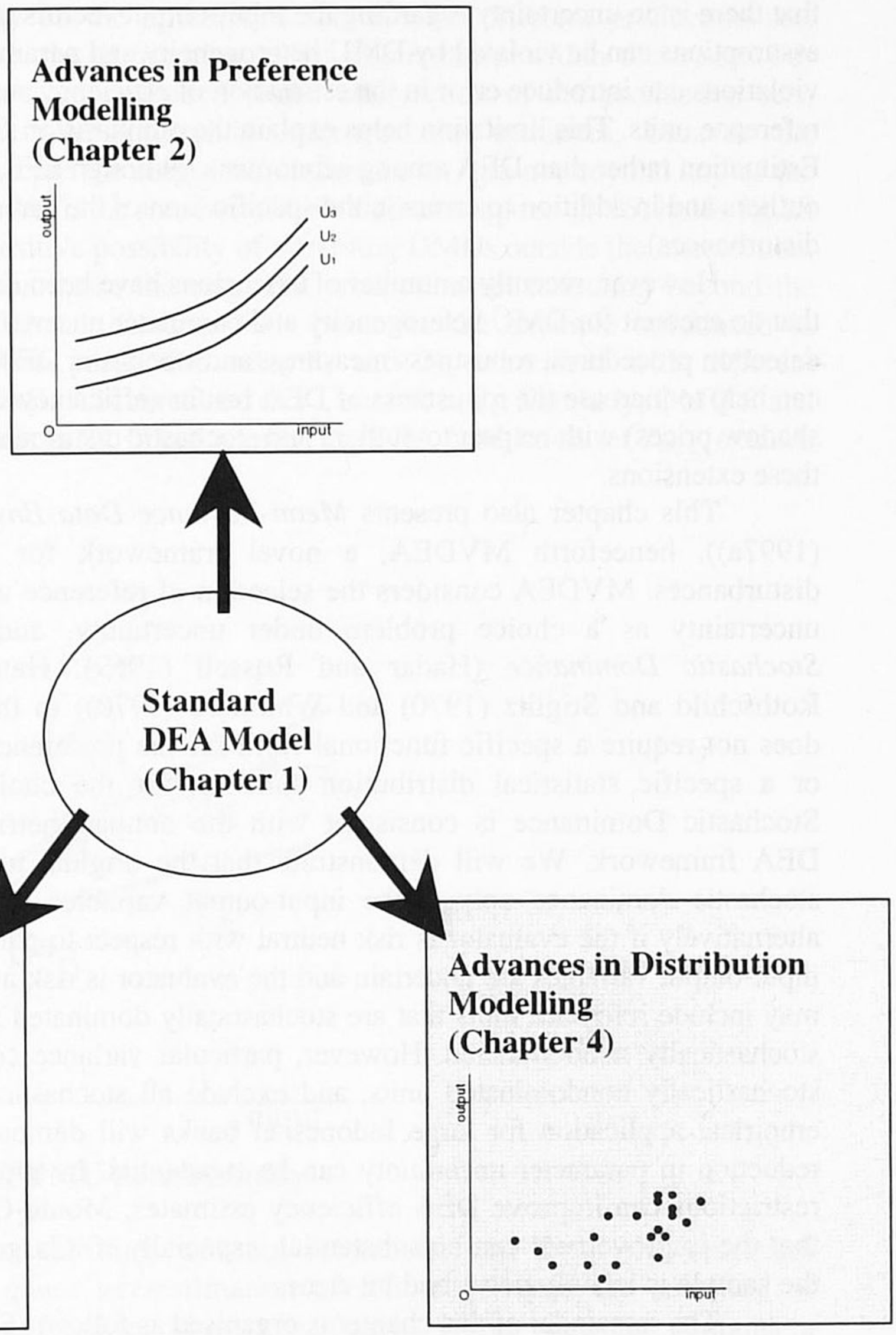
endo;
BIAS[T,4]=(sumc((Estimators[.,4]-Inefficiency))/SampleSize);
   STD[T,4]=stdc(Estimators[.,4]-Inefficiency);
   RMSE[T,4]=sqrt(BIAS[T,4]^2+STD[T,4]^2);

/******************** END TDEA-T model ************************/


   T=T+1;
   endo;

retp(BIAS,STD,RMSE);
endp;
```

Chapter 4:
Advances in Distribution Modelling



4.1. Introduction

As discussed in Chapter 1, the DEA methodology does not require a specific statistical distribution function for the inefficiencies. However, the original DEA models (including the BCC model discussed in Chapter 1 and used as a benchmark throughout this study) do impose a number of distribution assumptions that can be restrictive in particular research environments. Specifically, the original models assume that the DMUs in the data set represent a sample from a common production set, and in addition that there is no uncertainty regarding the input-output vectors of the DMUs. These assumptions can be violated by DMU heterogeneity and parameter uncertainty. Such violations can introduce error in the estimation of efficiency, and the selection of reference units. This limitation helps explain the emphasis on Econometric Frontier Estimation rather than DEA among economists (although EFE is also sensitive to outliers and in addition to errors in the specification of the statistical distribution of the disturbances).

However, recently a number of extensions have been developed in the literature that do account for DMU heterogeneity and parameter uncertainty, including outlier detection procedures, robustness measures, and stochastic DEA models. These advances can help to increase the robustness of DEA results (efficiency scores, reference units and shadow prices) with respect to outliers and stochastic disturbances. This chapter reviews these extensions.

This chapter also presents *Mean-Variance Data Envelopment Analysis* (Post (1997a)), henceforth MVDEA, a novel framework for incorporating stochastic disturbances. MVDEA considers the selection of reference units in case of parameter uncertainty as a choice problem under uncertainty, and applies the theory of *Stochastic Dominance* (Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970) and Whitmore (1970)) to that problem. That theory does not require a specific functional form for the preferences of the decision-maker or a specific statistical distribution function for the choice variables. Therefore, Stochastic Dominance is consistent with the nonparametric nature of the original DEA framework. We will demonstrate that the original models are consistent with stochastic dominance only if the input-output variables are specified accurately or alternatively if the evaluator is risk neutral with respect to parameter uncertainty. If the input-output variables are uncertain and the evaluator is risk averse, the original models may include reference units that are stochastically dominated and exclude units that are stochastically nondominated. However, particular variance corrections can include all stochastically nondominated units, and exclude all stochastically dominated units. An empirical application for large Indonesian banks will demonstrate that the associated reduction in parameter uncertainty can be substantial. In addition, particular variance restrictions can improve DEA efficiency estimates. Monte-Carlo simulations suggest that the improvement can be substantial, especially if a large sample is available and the sample is heavily perturbed by error.

The remainder of this chapter is organised as follows. Section 4.2 reviews the advances in the literature for obtaining results that are robust with respect to outliers and stochastic disturbances. Section 4.3 discusses the concept of stochastic dominance and introduces the proposed mean-variance framework. Section 4.4 discusses some special cases that reduce the information requirement and the computational burden. Section 4.5 illustrates the basic concepts and characteristics of the mean-variance approach by means of an empirical application for a real-life data set of large Indonesian banks. Section 4.6 assesses the potential for improving efficiency estimates using Monte-Carlo simulations. Finally, Section 4.7 offers some concluding remarks.

4.2. Improving DEA robustness

DMU heterogeneity occurs if the DMUs face different production sets. For example, a bank branch that operates in a metropolitan area can face a demographic and competitive situation that differs substantially from that of a branch in a rural area. The influence of heterogeneity on the goodness of DEA estimates and reference units is difficult to assess, because it depends heavily on the differences between the different production sets involved and in addition on the distribution of the observations within these sets. For example, if all observations are contained in the intersection of the production sets, introducing heterogeneous DMUs actually improves the estimates, because the additional observations increase the empirical production set, but not beyond any of the true production sets. However, if the standard distribution assumption (1.15) holds for all DMUs, there is a strictly positive possibility of observing DMUs outside the intersection. Therefore, heterogeneity can cause the empirical production set to extend beyond the true production sets. Figure 4.1 illustrates this point for the example introduced in Section 1.3. It displays two production sets, T_1 and T_2, associated with different production technologies. DMU A is contained in T_1, but not in T_2. Similarly, DMU F and G are contained in T_2, but not in T_1. Consequently, the DEA production set T_{BCC}, extends beyond both T_1 and T_2.

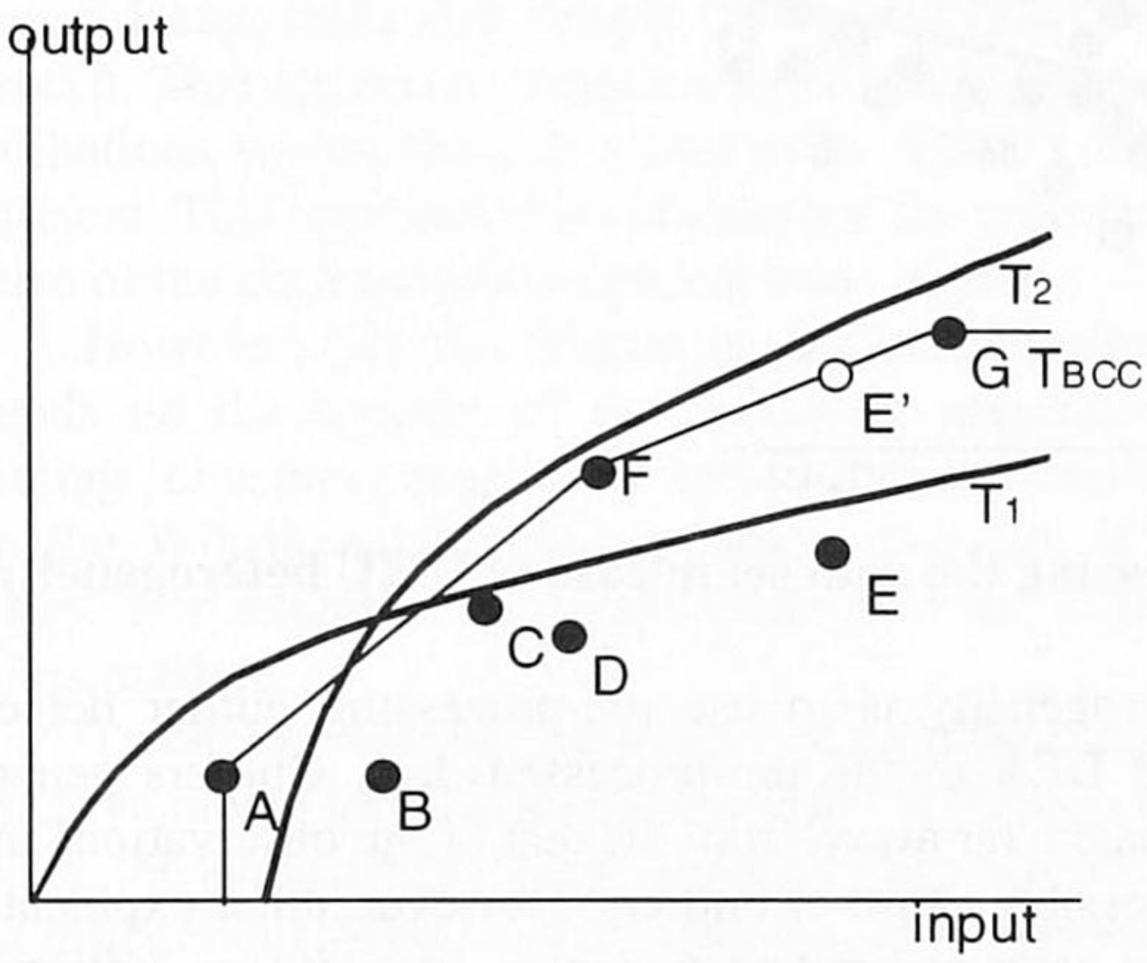


Figure 4. 1 The effect of DMU heterogeneity

Since DEA is based on selecting extreme reference units from the empirical production set, this can cause overestimation of efficiency[59] and the selection of reference units that are infeasible. More specifically, if the data set contains a reference unit that consumes less input than the evaluated unit and produces more than the efficient amount of output, the estimator exceeds true efficiency and the reference unit is not feasible, i.e.:

---

[59] Remember that efficiency is measured here as the maximum equiproportionate output augmentation, i.e. Debreu-Farrell output efficiency (1.3). The higher this measure the less efficient a DMU is. Therefore 'overestimation of efficiency' refers to the DMU appearing less efficient than it actually is. By contrast, if efficiency is measured as the direct output distance measure by Shephard (1953, 1970), i.e. the inverse of Debreu-Farrell output efficiency, a higher value is associated with more efficiency.

(4. 1) $\exists\lambda : X\lambda \le x_k; Y\lambda > \theta_k y_k; \lambda^T e = 1; \lambda \ge 0 \Rightarrow \hat{\theta}_{BCC,k} > \theta_k, (Y\lambda, X\lambda) \notin T$.

For example, in figure 4.1, DMU E is compared with a convex combination of DMU F and DMU G (say DMU E'). Both DMU F and DMU G operate at production levels which are not feasible for DMU E.

Increasing the sample can not eliminate the heterogeneous DMUs, and in fact introduces the possibility of introducing additional heterogeneous DMUs. Therefore, heterogeneity generally interferes with the statistical consistency of the DEA estimators. Figure 4.2 illustrates the effect of introducing additional observations in Figure 4.1. The additional observations bring the empirical production set closer to the union of both production sets, which overestimates the production possibilities for both technologies.

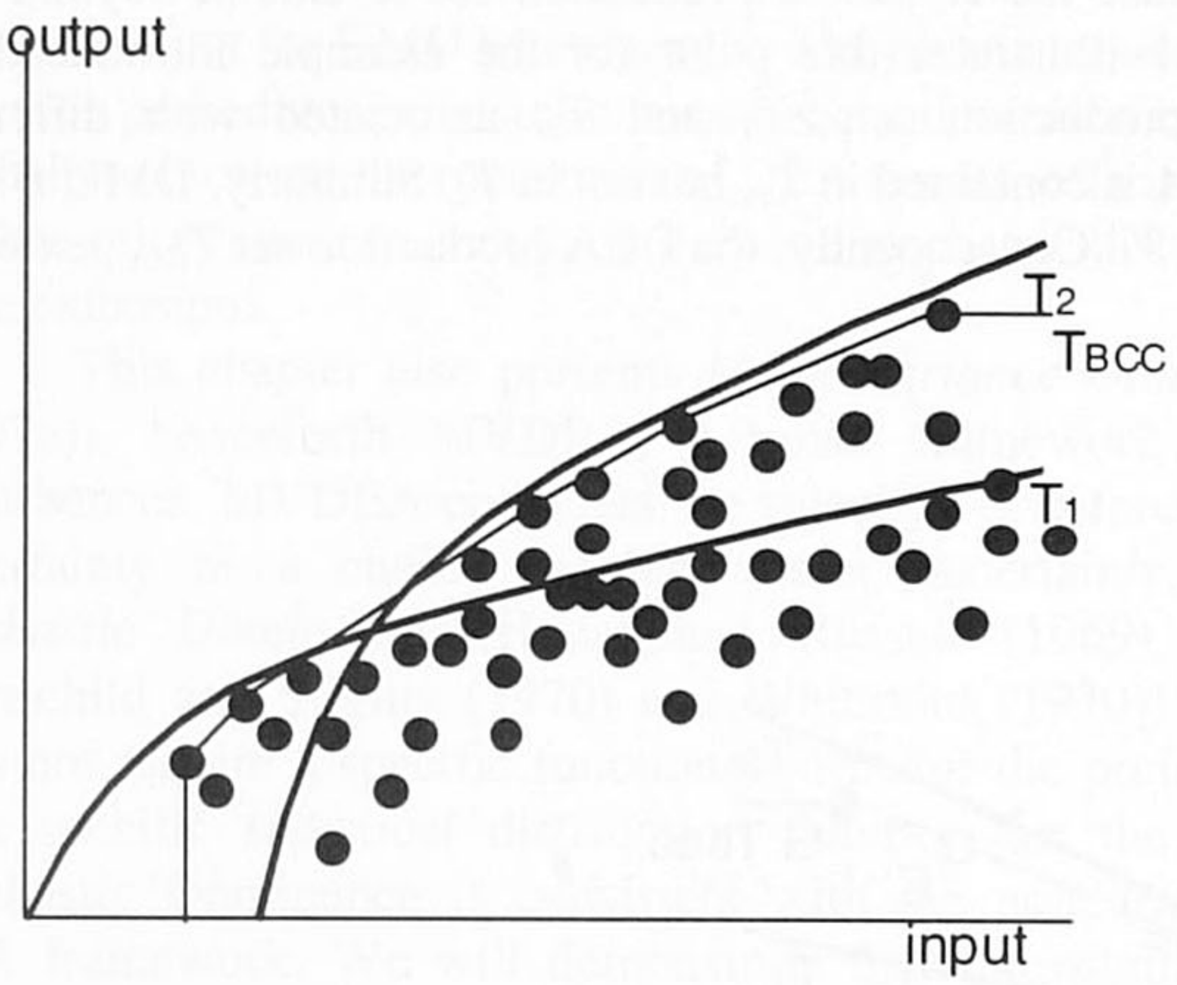


Figure 4. 2 The effect of increasing the data set in case of DMU heterogeneity

One way of dealing with heterogeneity is to use pre-processing outlier detection routines, and to apply standard DEA to the pre-processed data. Outliers generally refer to observations that are located 'far away' from the rest of the observations in the data set. Heterogeneity is one possible cause of outliers. However, other explanations may exist, such as measurement error or omitted variables. In addition, outliers can represent the most interesting observations in the data set, and excluding them can substantially reduce the information content of the sample. Therefore, outlier detection procedures are generally used for prioritising a follow-up inspection of the data. Unfortunately, as pointed out by Wilson (1995), standard methods for detecting outliers in regression models can not be adepted to DEA, because DEA essentially has a deterministic structure, since it assumes that there is no uncertainty regarding the input-output vectors of the DMUs.

Wilson (1995) proposed one of the sparse outlier detection routines specially tailored to DEA. That procedure relies on assessing the impact of excluding observations from the data set. If the exclusion of a particular DMU has a large impact on the efficiency scores of the remaining DMUs, the input-output vector of that DMU is supported by few additional observations. Consequently, the observation is a potential outlier and it is assigned a high priority for follow-up inspection. A

careful follow-up inspection of the data could reveal whether the observation has to be adjusted, omitted or can be included. This approach can be very useful if data checking is costly and resources are scarce. However, the model relies on the ability of the analyst to identify the outliers from the set of prioritized observations. In addition, the modified model excludes the evaluated unit only. This approach may fail if the data set contains multiple outliers. For example, the omission of an outlier can have little impact if one or more other outliers mask it in input-output space.

Alternatively, outliers could be detected using techniques that measure the robustness of DEA efficient classifications. As Wilson (1995) points out, it is important not to confuse sensitivity analysis with outlier detection. Sensitivity analysis checks the robustness of the results with respect to deviations of DMUs from their initial location in input-output space. By contrast, outlier detection involves finding DMUs whose location is atypical. Still, sensitivity analysis can be used for outlier detection purposes. Specifically, DMUs with a very robust efficient classification are potential outliers, because the data set contains little additional empirical evidence to support the attainability of their input-output combination.

A number of techniques have been proposed for sensitivity analysis in DEA, including Charnes, Cooper, Lewin and Morey (1985) and Charnes and Neralic (1990). Unfortunately, these studies hinge on conditions that preserve the feasibility and the optimality conditions of the optimal basis solution obtained via an extremal point algorithm. These conditions include non-degeneracy of the optimal basis and uniqueness of the optimal solution. These conditions are generally violated in DEA[60]. Charnes, Haag, Jaska and Semple (1992) and Zhu (1996) offer a more comprehensive approach. That approach computes a particular 'region of stability', a cell such that all perturbations within the cell preserve the DMU's current classification- efficient or inefficient. This approach does not require the calculation of the basis solution and the inverse of the corresponding optimal basis matrix.

However, like the Wilson procedure, the effectiveness of sensitivity analysis depends on the success of the follow-up inspection. In addition, the models for assessing robustness essentially consider data variations for the evaluated unit only. Like the Wilson model, this approach may fail if the data set contains multiple outliers. For example, the classification of an outlier can be non-robust if other outliers mask it.

To circumvent these problems, Kuosmanen and Post (1999a) proposed an alternative model for estimating efficiency if the data set is contaminated with outliers, the Robust Efficiency Model (REM). The REM relies on excluding the set of most influential DMUs from the data set. That operation suffices to correct for the influence of outliers, without having to identify the outliers prior to the analysis. The proper number of exclusions can be selected using an empirical test procedure.

Another problem is *parameter uncertainty* for the input-output vectors of the DMUs in the data set[61]. Traditional DEA neglects parameter uncertainty by treating the observations as if they were the true values for the input-output variables to select reference units for efficiency estimation and performance benchmarking. Stochastic

---

[60] For example, both primal and dual model formulations can have multiple solutions. See Section 1.8 and 3.3.

[61] Bawa and Klein (1976, 1977) used the term *parameter uncertainty* to refer to the error associated with estimating the mean vector and the covariance matrix in portfolio optimisation problems. This problem is related to the problem of *errors in variables* in regression analysis, which arises if the independent variables are subject to estimation error.

disturbances such as measurement error, random noise, outlier observations or external effects may violate this assumption. In stochastic environments, the input-output variables are unknown and have to be estimated using prior theory or available economic data. In many cases, theory and data allow for an inaccurate estimation only, and there is uncertainty regarding the true input-output variables. For example, much empirical research uses accounting data that can give a flawed representation of the underlying economic values, e.g. because of arbitrary valuation and depreciation schemes. Figure 4.3 displays hypothetical confidence areas for the input-output vectors of the DMUs in the example. With a certain probability, the true values of the input-output vector of the DMUs take a value from this area. The confidence area for DMU B is relatively small, which reflects the relative accuracy with which the input-output vector is estimated. By contrast, the size of the confidence area of DMU *C* reflects inaccurate estimation.

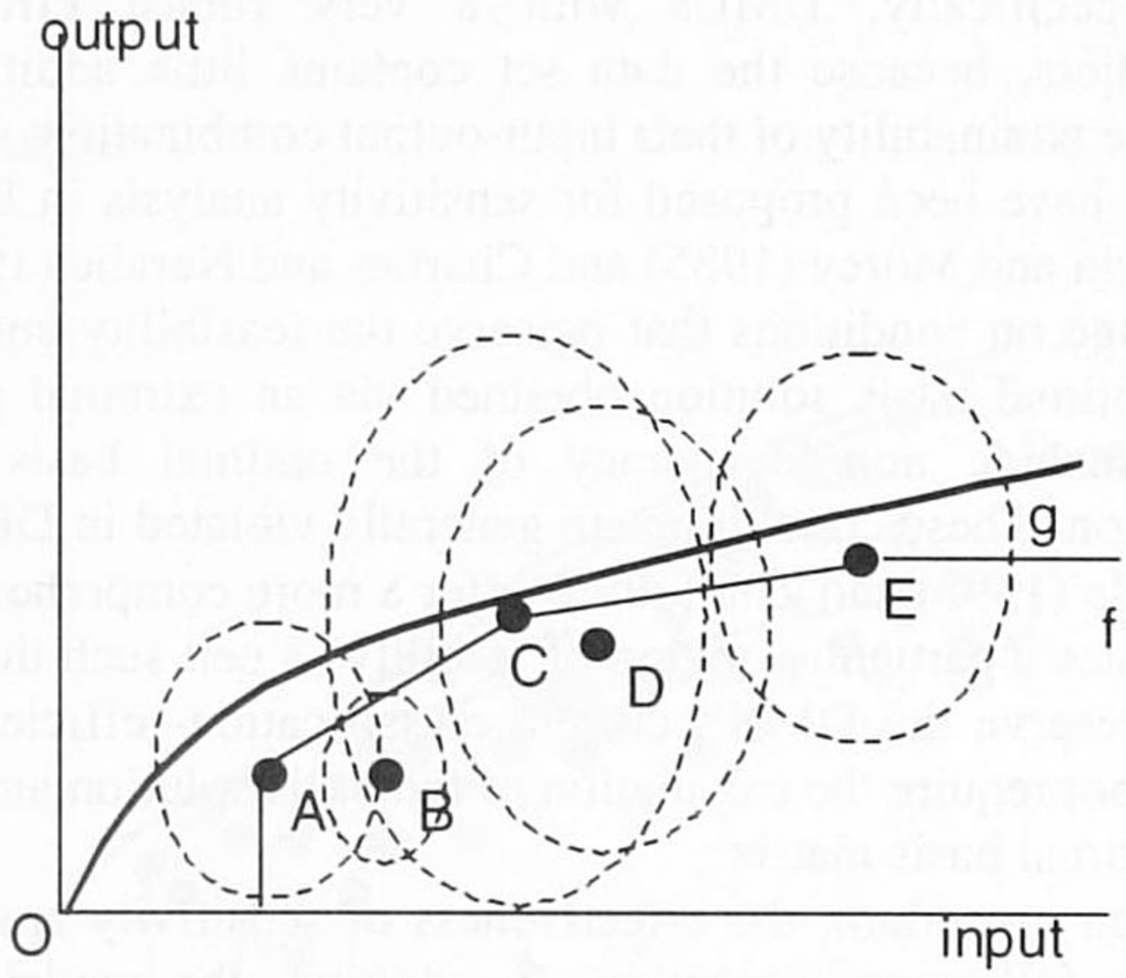


Figure 4. 3 the effect of parameter uncertainty

Parameter uncertainty complicates efficiency estimation and performance benchmarking, because it causes uncertainty for the position of the input-output vectors of DMUs relative to that of other DMUs, and relative to the production possibilities. The influence of parameter uncertainty on the goodness of DEA estimators and reference units is difficult to assess in general, because it depends heavily on the distribution of the observations and the disturbances. However, uncertainty for the input-output levels generally causes uncertainty for the efficiency measure, as the evaluated input-output vector becomes uncertain and in addition the empirical production set becomes uncertain and can be biased above the true production set. Parameter uncertainty represents a source of error in addition to sampling variation (see Section 1.4). However, unlike sampling variation, increasing the sample size can not eliminate this source of error. Figure 4.4 Figureically demonstrates the effect of increasing the data set in case of parameter uncertainty. The additional observations do not correct for the measurement error of the original observations, and in addition introduce additonal error.

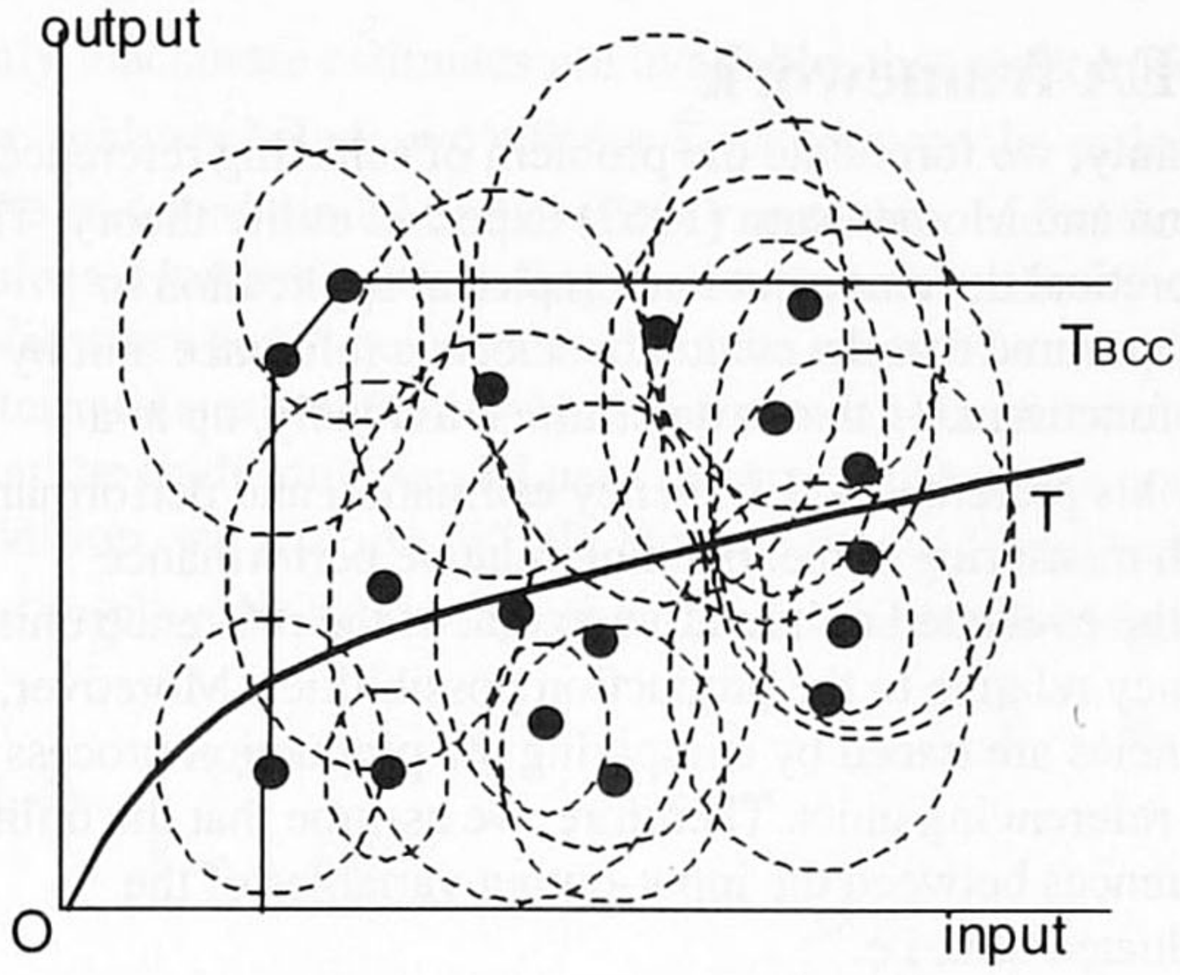


Figure 4. 4 The effect of increasing the data set in case of parameter uncertainty

Parameter uncertainty also complicates the selection of reference units. The referencing input-output values are uncertain, both relative to the production possibilities and relative to the evaluated unit. In addition, since the reference unit is selected using an extreme point technique, the referencing input-output values generally give an overly optimistic impression of the performance of the reference unit, and can exceed the production possibilities.

Outlier detection routines and robustness analysis are unlikely to solve the problems in case of parameter uncertainty. These techniques can work well in case of heterogeneity, because excluding the outliers restores the original assumptions. However, they are concerned with partial perturbations of selected data entries only, and are aimed at eliminating a subset of DMUs, and hence can not deal with situations in which all observations are perturbed.

However, a number of stochastic DEA models have been developed that explicitly acknowledge the stochastic nature of the data. For example, using certainty equivalence of stochastic input-output variables (Gong and Sun (1995)), or chance constrained programming (Land *et al.* (1994), Olesen and Petersen (1995), Cooper *et al.* (1996), (1998), Li (1998)) can incorporate disturbances. Alternatively, noise can be filtered from DEA inefficiencies by estimating a parametric error structure, as in the so-called DEA+ approach (Gstach (1998)).

These techniques can increase the robustness of the DEA model with respect to parameter uncertainty, and hence improve the reliability of DEA results. However, these techniques invariably require the specification of a particular statistical distribution function for the disturbances. Unfortunately, that requirement is not consistent with the nonparametric orientation of the original DEA models. By contrast, the Mean-Variance Data Envelopment Analysis (MVDEA, Post (1997a)) approach below uses stochastic dominance rules that do not require a particular distribution function.

4.3. The general MV-DEA framework

To incorporate parameter uncertainty, we formulate the problem of selecting reference units in terms of the Von Neumann and Morgenstern (1967) expected utility theory. That theory has seen considerable theoretical development and empirical application in economics. More specifically, we assume that the evaluator selects a reference unit by maximising the value of a utility function $U(\cdot)$ that is determined uniquely, up to a positive linear transformation, by his preferences. Efficiency estimation and performance benchmarking are concerned with measuring and explaining relative performance differences. The performance of the evaluated unit relative to that of the reference unit is used as an estimate for the effciency relative to the production possibilities. Moreover, causes and remedies for inefficiencies are traced by comparing the production process of the evaluated unit and that of the referencing units. Therefore, we assume that the utility function is defined over the differences between the input-output variables of the reference unit and that of the evaluated unit, i.e.[62]:

$$\textbf{(4. 2)} \quad U(\cdot) = U(Y\lambda - y_k, X\lambda - x_k).$$

In addition, we adhere to the standard assumption that reference units are restricted to the set of combinations of observations that are technically attainable under the assumptions imposed upon the production technology. In this study, we impose the standard assumptions of homogeneity (1.5), free disposability (1.6), and convexity (1.7). Consequently, the reference units are restricted to the set of all convex combinations, i.e . A_{BCC} (1.17).Under these assumptions, the problem of selecting a reference unit can be represented as:

$$\textbf{(4. 3)} \ \underset{\lambda}{\text{Max}} \left\{ U(Y\lambda - y_k, X\lambda - x_k) \middle| \lambda \in A_{BCC} \right\}.$$

This formulation may leave out important complications faced in actual performance evaluation problems, such as the number and the identity of the benchmark partners[63]. The inclusion of such complications would require a refinement to the analysis provided in this study. However, as will be demonstrated below, under the assumed structure traditional DEA is consistent with expected utility maximization if the observations are not perturbed or alternatively if the evaluator is risk neutral with respect to parameter uncertainty. Consequently, the traditional framework can serve as a benchmark for environments involving disturbances and risk aversion.

To solve decision problem (4.3), the true input-output variables have to be specified. In contrast to the original models, we explicitly assume that the input-output variables are perturbed and cannot be specified accurately. We assume that

---

[62] The utility function introduced in Chapter 2 was defined over the inputs and outputs in absolute values. By contrast, the utility function used in this chapter is defined over input and output differences relative to the reference unit. For the deterministic case, the two approaches can be demonstrated to give identical results. However, for the stochastic case, the results can be different. A benefit of using performance differences is that it excludes infeasible solutions if variance restrictions are imposed. If the evaluated unit is compared with itself (which is a feasible solution, because the evaluated DMU is included in the data set), the variance of the performance improvements is zero by definition. Hence, variance restrictions, no matter how strong, can not cause infeasible solutions.

[63] For a discussion of the factors that determine success or failure in performance benchmarking, see Rolstaadas *et al.* (1995) and Spendolini (1992).

only inaccurate estimates are available, that randomly deviate from the true values. In the analysis below, we will let $\hat{Y}$ ($\hat{X}$) denote the estimates for the outputs (inputs). The standard model (1.10) simply uses the observations as estimates for the true values. However, depending on the particular theory and data available, alternative estimators could be used, such as the mean of a sample of multiple observations. If alternative estimates are used in the mean-variance model, we assume for simplicity that the traditional model uses these estimates also (rather than the observations). In addition, we will use W' (W) for the output (input) estimation errors. More formally, we employ the following structure for the input-output estimates:

$$\textbf{(4. 4)} \qquad \begin{aligned} \hat{Y} &= Y + W'; \\ \hat{X} &= X + W. \end{aligned}$$

In these expressions, $\hat{Y}$ and W' are (p x n) matrices and $\hat{X}$ and W are (m x n) matrices. In addition, $w'_{.k}$ ($w_{.k}$) and $\hat{y}_{.k}$ ($\hat{x}_{.k}$) denotes the k-th column of W' (W) and $\hat{Y}$ ($\hat{X}$) respectively. By analogy, in the analysis below, we will use $w'_{r.}$ ($w_{i.}$) and $\hat{y}_{r.}$ ($\hat{x}_{i.}$) to denote the r-th (i-th) row of W' (W) and $\hat{Y}$ ($\hat{X}$) respectively.

If these estimates are used rather than the true values for the input-output variables, selecting a reference unit is a problem of choice under uncertainty, as it involves parameter uncertainty. Under the above assumptions, it is consistent with the Von Neumann and Morgenstern (1967) axioms for the optimal reference unit to be given by the solution to the following maximisation problem:

$$\textbf{(4. 5)} \qquad \underset{\lambda}{\text{Max}} \left\{ E[U(((\hat{Y} - W')\lambda - (\hat{y}_{.k} - w'_{.k})), ((\hat{X} - W)\lambda - (\hat{x}_{.k} - w_{.k})))] \middle| \lambda \in A_{BCC} \right\}$$

$$= \underset{\lambda}{\text{Max}} \left\{ \int_{D(W',W)} U(((\hat{Y} - W')\lambda - (\hat{y}_{.k} - w'_{.k})), ((\hat{X} - W)\lambda - (\hat{x}_{.k} - w_{.k})))\partial F(W',W) \middle| \lambda \in A_{BCC} \right\}.$$

In these expressions, $F(W',W)$ and $D(W',W)$ denote the multivariate distribution function and the domain for the estimation errors respectively.

As is true for many problems of choice under uncertainty, this problem cannot be solved without making further assumptions regarding the probability distribution of the estimation errors and the preference structure of the evaluator. However, we intend to impose minimal preference and distribution assumptions, so as to minimize the information requirement and preserve the conservative nature of the conventional DEA framework. Therefore, we propose to use decision criteria from the theory of stochastic dominance (Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970) and Whitmore (1970)). That is a theory of choice under uncertainty that has seen considerable theoretical development and empirical application in economics in the last decades[64]. The most general form of stochastic dominance (henceforth SD) makes no assumptions about the form of the disturbance distribution. Furthermore, no specific functional form for decision-maker preferences is assumed.

---

[64] For a detailed survey and analysis of the expected utility and stochastic dominance literature, see Levy (1992).

Rather, SD relies on assumptions about the general characteristics of decision-maker preferences. There are three progressively stronger preference assumptions that are employed in the SD literature. They lead directly to the first-, second-, and third-order SD criteria. All three criteria assume that decision-makers are attempting to maximize the expected utility associated with the outcomes of their choices. First-order SD assumes decision-makers are non-satiable, i.e. they prefer more over less. Second-order SD (henceforth SSD) assumes, in addition to non-satiation, that decision-maker are risk averse, that is, they dislike uncertainty. Finally, third-order SD (henceforth TSD) adds to non-satiation and risk aversion the assumption that decision-makers have decreasing absolute risk aversion (henceforth DARA, Pratt (1964)), i.e. their dislike for absolute uncertainties decreases as the levels of the outcomes increase. Higher order criteria involve more discriminating power than lower order ones, because they induce a larger reduction of the set of nondominated choice alternatives. However, that power has to be balanced against the stringency of the preference assumptions imposed. In general, striking that balance requires a careful consideration of the structure and the context of the decision problem considered.

For the performance evaluation problem, assuming non-satiation corresponds to assuming that utility is monotonously increasing in the outputs and monotonously decreasing in the inputs. This seems a valid assumption, since the higher the outputs and the lower the inputs, the closer it is to the frontier and the higher the potential performance improvements that are revealed. Moreover, as discussed in the previous Section, nonsatiation is implicitly assumed in the standard model. In addition, we assume that the evaluator is risk averse, that is, he dislikes uncertainty for the input-output variables. In our opinion, this assumption is tenable, e.g. because the reward of the evaluating manager is tied to the performance improvements of the evaluated units, and moreover the evaluating manager has risk aversion for his personal wealth. In addition, DARA may be a tenable assumption, leading to the most powerful TSD criterion[65]. However, in the SD literature there exist no simple operational algorithms for fitting this criterion to empirical discrete data. Therefore, we propose using the less powerful SSD. That decision rule can easily be operationalised using general distribution assumptions.

In general, a random variable stochastically dominates another one by second order, if and only if, for any chosen value, the latter has more cumulative probability mass than the former. More formally, letting $F_A(\cdot)$ denote the cumulative probability density for random variable A and $F_B(\cdot)$ for B, A stochastically dominates B by second-order, if and only if

$$\textbf{(4. 6)} \quad \int_{-\infty}^{c} F_B(z)\partial z \geq \int_{-\infty}^{c} F_A(z)\partial z \quad \forall c \in \Re.$$

If this condition is satisfied, all expected utility maximising non-satiable and risk averse decision-makers will prefer A to B.

Figures 4.5 and 4.6 illustrate the concept of SSD for two hypothetical DMUs that produce a single stochastic output. Figure 4.5 displays the probability density function for both DMUs, and Figure 4.6 displays the associated cumulative probality density ('probability mass') functions. DMU A stochastically dominates B by second order, because A has a larger probablity of producing at least an arbitrary output level than B.

---

[65] For example, empirical evidence based on survey data supports DARA for investor asset choices (Blume and Fried (1975), Cohn *et al.* (1975)).

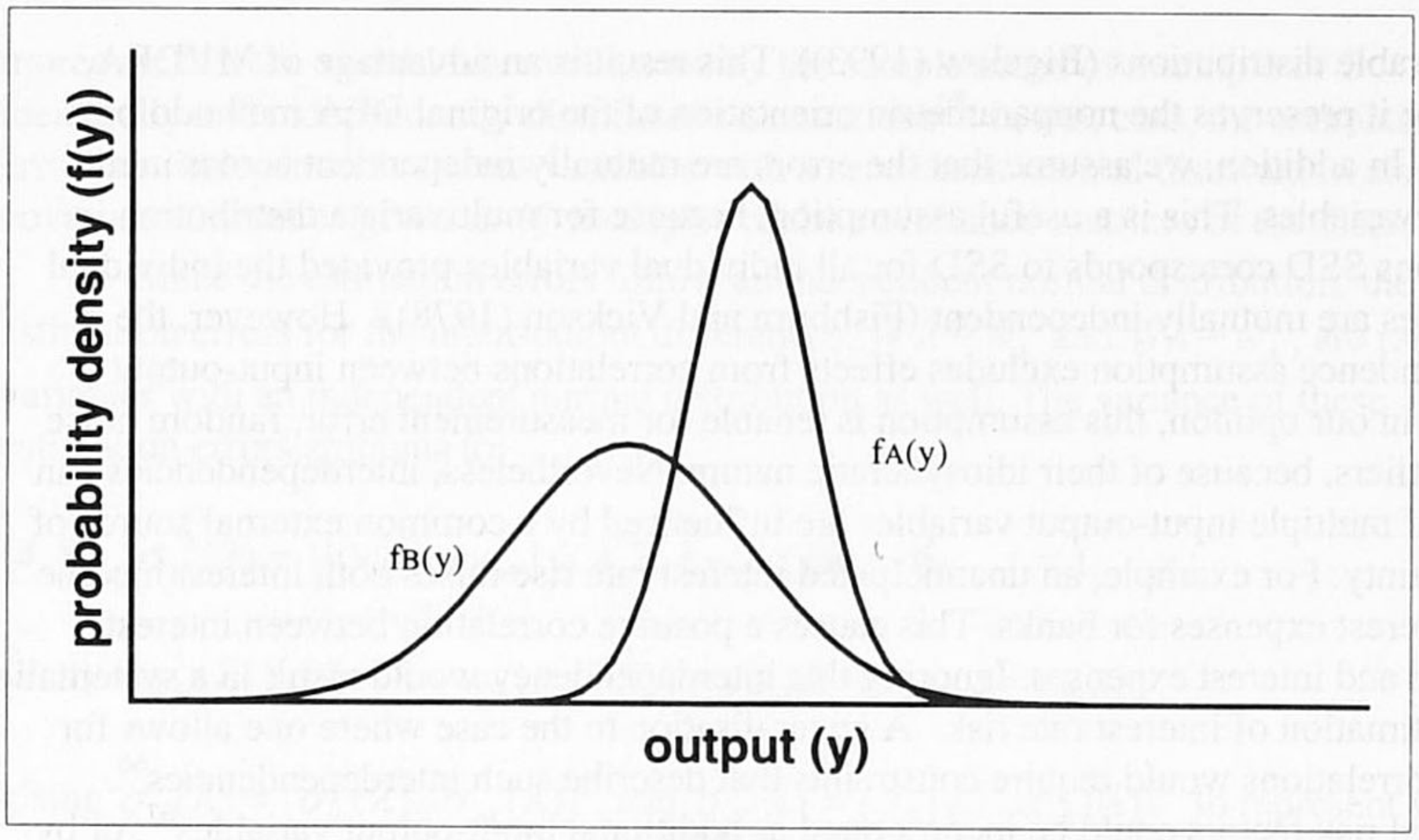


Figure 4. 5 The probability density functions of two random variables

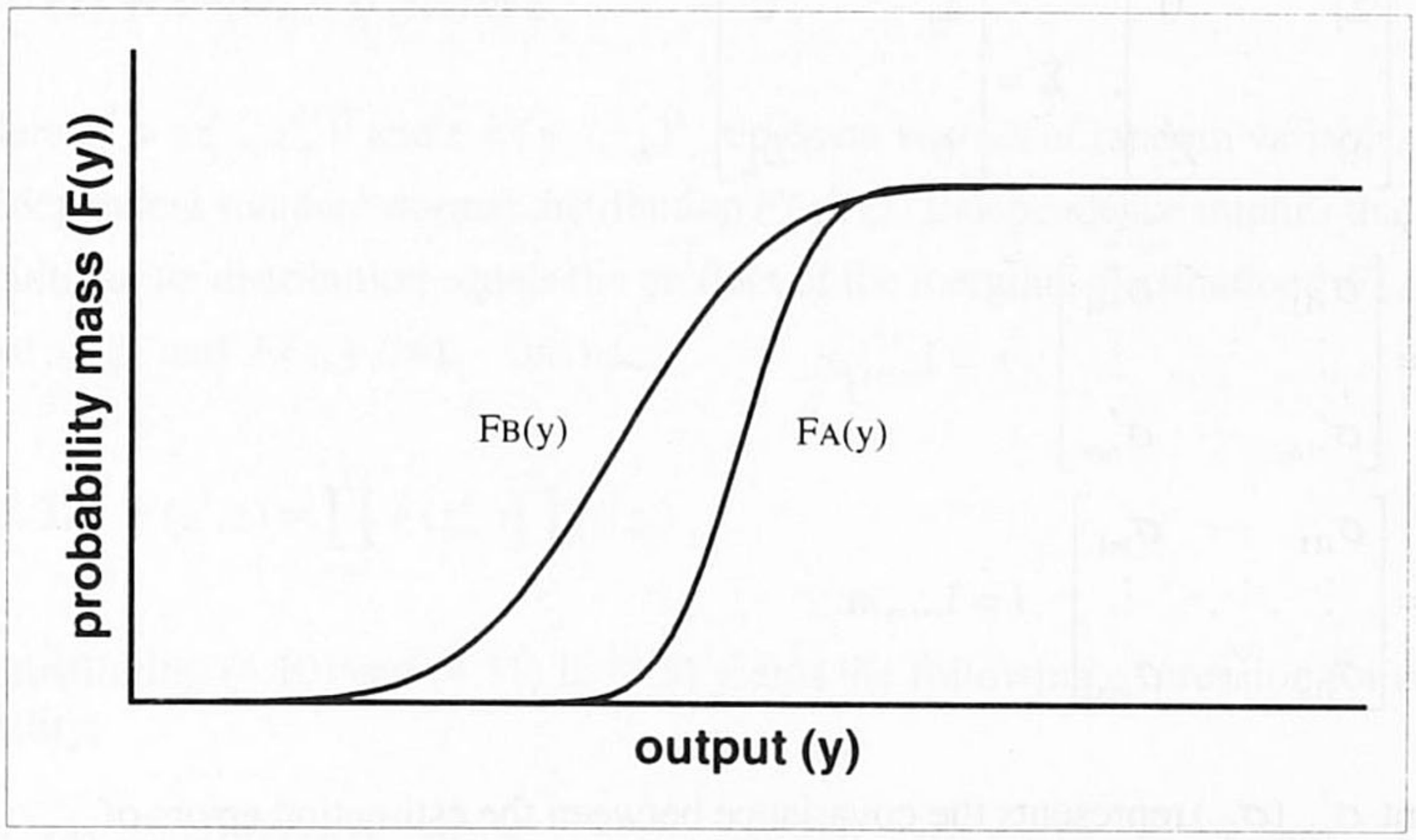


Figure 4. 6 The probability mass of the random variables

To operationalise the SSD criterion for selecting reference units, we have to further specify the distribution of the estimation errors. We assume here that the errors come from a multivariate normal distribution with zero means. This is a convenient assumption that helps to simplify the analysis, because for normally distributed random variables, SSD is equivalent to mean-variance dominance, i.e. a variable stochastically dominates another one by second order if it has a higher mean and a lower variance (Tobin (1958)). However, we stress that it is not necessary to impose a particular distribution function. Specifically, mean-variance analysis can be demonstrated to be consistent with expected utility theory for the much broader class of normalised risk-

comparable distributions (Bigelow (1993)). This result is an advantage of MVDEA, because it preserves the nonparametric orientation of the original DEA methodology.

In addition, we assume that the errors are mutually independent across input-output variables. This is a useful assumption, because for multivariate distribution functions SSD corresponds to SSD for all individual variables provided the individual variables are mutually independent (Fishburn and Vickson (1978)). However, the independence assumption excludes effects from correlations between input-output errors. In our opinion, this assumption is tenable for measurement error, random noise and outliers, because of their idiosyncratic nature. Nevertheless, interdependencies can occur if multiple input-output variables are influenced by a common external source of uncertainty. For example, an unanticipated interest rate rise raises both interest income and interest expenses for banks. This causes a positive correlation between interest income and interest expenses. Ignoring this interdependency would result in a systematic overestimation of interest rate risk. A generalisation to the case where one allows for such correlations would require constraints that describe such interdependencies[66]. External risk factors could be incorporated as additional input-output variables[67], or by correcting the input-output variables for systematic uncertainty, leaving idiosyncratic uncertainty only.

More formally, letting Σ' (Σ) denote the covariance matrix for the estimation errors for outputs (inputs), we assume the following covariance structure:

$$\textbf{(4. 7)} \quad \Sigma' = \begin{bmatrix} \Sigma'_1 & & 0 \\ & \ddots & \\ 0 & & \Sigma'_p \end{bmatrix}; \quad \Sigma = \begin{bmatrix} \Sigma_1 & & 0 \\ & \ddots & \\ 0 & & \Sigma_m \end{bmatrix};$$

$$\Sigma'_r = \begin{bmatrix} \sigma'_{r11} & \cdots & \sigma'_{rm1} \\ \vdots & \ddots & \vdots \\ \sigma'_{r1n} & \cdots & \sigma'_{rmn} \end{bmatrix} \quad r = 1,\ldots, p;$$

$$\Sigma_i = \begin{bmatrix} \sigma_{i11} & \cdots & \sigma_{in1} \\ \vdots & \ddots & \vdots \\ \sigma_{i1n} & \cdots & \sigma_{inn} \end{bmatrix} \quad i = 1,\ldots, m.$$

Here element σ'_{rlj} (σ_{ilj}) represents the covariance between the estimation errors of observation j and observation l for output r (input i). In the analysis below, we will use $\sigma'_{rl\cdot}$ ($\sigma_{il\cdot}$) to denote the l-th row of the covariance matrix for output r (input i).

Obviously, the validity of the above distribution assumptions (normality, zero means and independence) depends on the particular estimators used for the input-output variables. One way of justifying them is by assuming that the input-output variables are perturbed by disturbances from a distribution that satisfies these assumptions, and

---

[66] Post and Spronk (1999b) develop such a model based on multi-factor risk models.

[67] This would mean that a DMU is efficient if it were not dominated in the input-output space by composite units that are less favorably influenced by external risks. Whether these reference units can be constructed as positive linear combinations, convex combinations, non-composite units, or in some other way, depends on the assumptions the analyst imposes on the nature of the external risks.

moreover that the variables are estimated by the sample mean of a sample of T identically and independently distributed observations[68]. In that case, the estimation errors are independent random variables with a zero-mean normal distribution also, with covariance matrices given as $1/T$ multiples of the covariance matrices of the disturbances.

Since the estimation errors follow an independent normal distribution, the estimation errors for the input-output differences, $W'\lambda - w'_{\cdot k}$ and $W\lambda - w_{\cdot k}$, are random variables with an independent normal distribution as well. The variance of these estimation errors amounts to:

(4. 8) $\sigma'^2_{rk}(\lambda) = V[w'_{r\cdot}\lambda - w'_{rk}] = \lambda^T \Sigma'_r \lambda - 2\sigma'_{rk\cdot}\lambda + \sigma'_{rkk} \quad r = 1,\ldots, p$;

(4. 9) $\sigma^2_{ik}(\lambda) = V[w_{i\cdot}\lambda - w_{ik}] = \lambda^T \Sigma_i \lambda - 2\sigma_{ik\cdot}\lambda + \sigma_{ikk} \quad i = 1,\ldots, m$.

Using $\sigma'_{\cdot k}(\lambda) = (\sigma'_{ik}(\lambda)\ldots\sigma'_{pk}(\lambda))^T$ and $\sigma_{\cdot k}(\lambda) = (\sigma_{ik}(\lambda)\ldots\sigma_{mk}(\lambda))^T$ to represent vectors of the corresponding standard deviations, and $\otimes$ for the element-by-element multiplication of vectors, the estimation errors can be expressed as:

(4. 10) $W'\lambda - w'_{\cdot k} = \sigma'_{\cdot k}(\lambda) \otimes z'$;

(4. 11) $W\lambda - w_{\cdot k} = \sigma_{\cdot k}(\lambda) \otimes z$.

Here $z' = (z'_1 \ldots z'_p)^T$ and $z = (z_1 \ldots z_m)^T$ represent vectors of random variables with an independent standard normal distribution $F(z', z)$. Independence implies that the multivariate distribution equals the product of the marginal distributions $F(z'_r)$, r=1,…,p, and $F(z_i)$, i=1,…,m, i.e.:

(4. 12) $F(z', z) = \prod_{r=1}^{p} F(z'_r) \prod_{i=1}^{m} F(z_i)$.

Substituting (4.10) and (4.11) in (4.5) yields the following expression for expected utility:

(4. 13) $\max_{\lambda} \left\{ E[U(((\hat{Y} - W')\lambda - (\hat{y}_{\cdot k} - w'_{\cdot k})), ((\hat{X} - W)\lambda - (\hat{x}_{\cdot k} - w_{\cdot k})))] \middle| \lambda \in A_{BCC} \right\} =$

$$\max_{\lambda} \left\{ \int_{-\infty}^{\infty} U((\hat{Y}\lambda - \hat{y}_{\cdot k} + \sigma'_{\cdot k}(\lambda) \otimes z'), (\hat{X}\lambda - \hat{x}_{\cdot k} + \sigma_{\cdot k}(\lambda) \otimes z)) \prod_{r=1}^{p} \partial F(z'_r) \prod_{i=1}^{m} \partial F(z_i) \middle| \lambda \in A_{BCC} \right\}.$$

In this expression, expected utility depends on the estimated input-output differences and the error variances only. Therefore, the decision problem can be represented by:

(4. 14) $\underset{\lambda}{\text{Max}} \left\{ E[U(\hat{Y}\lambda - \hat{y}_{\cdot k}, \hat{X}\lambda - \hat{x}_{\cdot k}, \sigma'^2_{\cdot k}(\lambda), \sigma^2_{\cdot k}(\lambda))] \middle| \lambda \in A_{BCC} \right\}$.

---

[68] This includes the traditional approach of using observations as estimates. In that case, the estimation errors coincide with the disturbances.

This differs from the traditional framework, because it involves the error variances as goal variables in addition to the estimated input-output differences.

Traditional DEA assumes that the evaluator prefers more over less for outputs and less over more for inputs. That assumption corresponds to the non-satiation assumption. As discussed above non-satiation seems a valid assumption for performance evaluation. However, in case of parameter uncertainty, it is generally not sufficient to restrict the reference units to the BCC admissible set (1.17), because the error variances have to be considered also. The traditional approach implicitly assumes that these terms can be ignored or equivalently that the observations are unperturbed or alternatively that the evaluator is risk neutral. By contrast, we assume in addition to non-satiation that the evaluator is risk averse. Risk aversion implies that expected utility is monotonously decreasing in the error variances. Therefore, the admissible set can be reduced to all composite units that are not dominated in mean-variance terms, i.e. no alternative units exist that give both higher (lower) output (input) estimates and lower variances. More formally, that set amounts to:

$$\textbf{(4. 15)}\ A_{MV} = \left\{ \lambda \in \Re_+^n \middle| \left\{ \tau \in \Re_+^n \middle| \begin{pmatrix} \hat{Y}\tau \\ -\hat{X}\tau \end{pmatrix} - \begin{pmatrix} s' \\ s \end{pmatrix} = \begin{pmatrix} \hat{Y}\lambda \\ -\hat{X}\lambda \end{pmatrix}; \begin{pmatrix} \sigma'_{.k}(\tau) \\ \sigma_{.k}(\tau) \end{pmatrix} + \begin{pmatrix} d' \\ d \end{pmatrix} = \begin{pmatrix} \sigma'_{.k}(\lambda) \\ \sigma_{.k}(\lambda) \end{pmatrix}; \right.\right.$$
$$\left.\left. \tau^T e = 1; s', d', s, d \geq 0; e'^T (s' + d') + e^T (s + d) > 0 \right\} = \varnothing; \lambda^T e = 1 \right\}.$$

In this expression $d' = (d'_1 \ldots d'_p)^T$ and $d = (d_1 \ldots d_m)^T$ represent the excess error variances, i.e. the differences between the error variances of the composite units characterized by intensity vectors τ and λ.

We will refer to this set as the mean-variance admissible set. As a result of parameter uncertainty and risk aversion, it differs from the admissible set obtained by traditional analysis. More specifically, the mean-variance set excludes all composite units that are nondominated in terms of input-output estimates but for which the data set contains alternative units with identical input-output estimates, but lower error variances, i.e.

$$\textbf{(4. 16)} \left\{ \lambda \in \Re_+^n \middle| \left\{ \tau \in \Re_+^n \middle| \begin{pmatrix} \hat{Y}\tau \\ -\hat{X}\tau \end{pmatrix} - \begin{pmatrix} s' \\ s \end{pmatrix} = \begin{pmatrix} \hat{Y}\lambda \\ -\hat{X}\lambda \end{pmatrix}; \tau^T e = 1; s', s \geq 0; e'^T s' + e^T s > 0 \right\} = \varnothing; \right.$$
$$\left\{ \tau \in \Re_+^n \middle| \begin{pmatrix} \hat{Y}\tau \\ -\hat{X}\tau \end{pmatrix} = \begin{pmatrix} \hat{Y}\lambda \\ -\hat{X}\lambda \end{pmatrix}; \begin{pmatrix} \sigma'_{.k}(\tau) \\ \sigma_{.k}(\tau) \end{pmatrix} + \begin{pmatrix} d' \\ d \end{pmatrix} = \begin{pmatrix} \sigma'_{.k}(\lambda) \\ \sigma_{.k}(\lambda) \end{pmatrix}; \right.$$
$$\left. \left. \tau^T e = 1; d', d \geq 0; e'^T d' + e^T d > 0 \right\} \neq \varnothing; \lambda^T e = 1 \right\}.$$

Risk neutral evaluators might select a reference unit from this set, because its elements are nondominated in terms of input-output estimates. However, for risk-averse evaluators this set can be discarded, because it is dominated in terms of error variance. In addition, in contrast to the BCC set, the mean-variance set includes all units that are dominated in terms of input-output estimates, but not in terms of error variance, i.e.:

$$\textbf{(4. 17)}\left\{\lambda\in\Re_+^n\middle|\left\{\tau\in\Re_+^n\middle|\begin{pmatrix}\hat{Y}\tau\\-\hat{X}\tau\end{pmatrix}-\begin{pmatrix}s'\\s\end{pmatrix}=\begin{pmatrix}\hat{Y}\lambda\\-\hat{X}\lambda\end{pmatrix};\ \tau^Te=1;s',s\geq 0;\ e'^Ts'+e^Ts>0\right\}\neq\varnothing;\right.$$

$$\left.\left\{\tau\in\Re_+^n\middle|\begin{pmatrix}\sigma'_{.k}(\tau)\\\sigma_{.k}(\tau)\end{pmatrix}+\begin{pmatrix}d'\\d\end{pmatrix}=\begin{pmatrix}\sigma'_{.k}(\lambda)\\\sigma_{.k}(\lambda)\end{pmatrix};\ \tau^Te=1;d',d\geq 0;\ e'^Td'+e^Td>0\right\}=\varnothing;\ \lambda^Te=1\right\}.$$

All non-satiable risk neutral evaluators discard these units, because the data set contains alternative units with better input-output estimates. However, risk-averse evaluators might select reference units from this set, because they are nondominated in terms of error variance.

Using mean-variance set (4.15) rather than BCC set (1.17), decision problem (4.3) can be represented as:

$$\textbf{(4. 18)}\ \max_{\lambda}\left\{E\left(U\left(\hat{Y}\lambda-\hat{y}_{.k},\hat{X}\lambda-\hat{x}_{.k},\sigma'_{.k}(\lambda),\sigma_{.k}(\lambda)\right)\right)\middle|\lambda\in A_{MV}\right\}.$$

As in traditional DEA, reference units can be selected from this set using a mathematical programming problem that reflects preference information in addition to non-satiation and risk-aversion, or alternatively using an interactive decision procedure. For example, adhering to the standard assumption that the evaluator prefers a reference unit that uses the lowest possible fraction of the inputs of the evaluated unit to produce at least its outputs, model (1.20) can be extended as follows:

$$\textbf{(4. 19)}\qquad \max_{\theta,\lambda_j,s_i,s'_r}\ \theta+\varepsilon\left[\sum_{i=1}^{m}s_i+\sum_{r=1}^{p}s'_r\right]$$

$$+\varepsilon\left[\sum_{i=1}^{m}\left(\sum_{j=1}^{n}\sum_{l=1}^{n}\lambda_j\lambda_l\sigma_{ilj}-2\sum_{j=1}^{n}\lambda_j\sigma_{ikj}+\sigma_{ikk}\right)+\sum_{r=1}^{p}\left(\sum_{j=1}^{n}\sum_{l=1}^{n}\lambda_j\lambda_l\sigma'_{rlj}-2\sum_{j=1}^{n}\lambda_j\sigma'_{rkj}+\sigma'_{rkk}\right)\right]$$

$$\text{s.t.}\ \sum_{j=1}^{n}\lambda_j\hat{y}_{rj}-s'_r=\theta\hat{y}_{rk}\qquad r=1,\ldots,p;$$

$$\sum_{j=1}^{n}\lambda_j\hat{x}_{ij}+s_i=\hat{x}_{ik}\quad i=1,\ldots,m;$$

$$\sum_{j=1}^{n}\lambda_j=1;$$

$$\lambda_j\geq 0\quad j=1,\ldots,n;$$

$$s'_r\geq 0\quad r=1,\ldots,p;$$

$$s_i\geq 0\quad i=1,\ldots,m.$$

The objective function of this model includes a 'non-Archimedean' variance correction in addition to the slack correction used in the traditional model. That variance correction is included to exclude the selection of a reference unit that is dominated by other units that give an identical radial input reduction, but have less error variance, i.e. an element of set (4.16). The units that are dominated in terms of input-output estimates but nondominated in terms of error variance, i.e. the elements

of set (4.17), can be discarded here, because the contraction of inputs and the production of DMU_k's outputs is given pre-emptive priority over variance reductions. Therefore, the variance correction effectively restricts the reference units to the mean-variance set. The standard model (1.20) can be solved by first maximizing the radial gain and subsequently maximizing a target sum to obtain further gains where possible, retaining all the radial gain. Most DEA-software (e.g. Warwick software[69]) embeds such a two-step procedure. A similar two-stage procedure can be used to solve (4.19). More specifically, the first stage is finding the optimal value for δ by solving the following problem:

$$\textbf{(4.20)} \quad \max_{\theta,\lambda_j,s_i,s_r'} \delta$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j \hat{y}_{rj} - s_r' = \theta \hat{y}_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j \hat{x}_{ij} + s_i = \hat{x}_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s_r' \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

The second stage uses the optimal solution, say θ^*, to find the optimal reference unit by solving the following problem:

$$\textbf{(4.21)} \quad \max_{\lambda_j,s_i,s_r'} \sum_{r=1}^{p}\left(s_r' + \sum_{j=1}^{n}\sum_{l=1}^{n} \lambda_j \lambda_l \sigma'_{rlj} - 2\sum_{j=1}^{n} \lambda_j \sigma'_{rkj} + \sigma'_{rkk} \right)$$

$$+ \sum_{i=1}^{m}\left(s_i + \sum_{j=1}^{n}\sum_{l=1}^{n} \lambda_j \lambda_l \sigma_{ilj} - 2\sum_{j=1}^{n} \lambda_j \sigma_{ikj} + \sigma_{ikk} \right)$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j \hat{y}_{rj} - s_r' = \hat{y}_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j \hat{x}_{ij} + s_i = \theta^* \hat{x}_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s_r' \geq 0 \quad r = 1,\ldots,p;$$

$$s_i \geq 0 \quad i = 1,\ldots,m.$$

---

[69] See Ali Emrouznejad's DEA homepage', http://www.csv.warwick.ac.uk/~bsrlu/dea/dea1.htm.

Example 4.1 To illustrate the effect of including a variance correction, consider 4 hypothetical DMUs with input-output estimates and standard deviations given in Table 4.1. Assume for simplicity that the estimation errors are independent across variables and across DMUs.

	$\hat{x}$	$\hat{y}$	σ_x	σ_y
A	1	2	1	1
B	2	1	1	1
C	3	3	2	2
D	5	4	0.5	0.5

Table 4. 1 Example data set

Figure 4.7 displays the input-output estimates. In addition, it displays confidence areas with diameters that are equal to the standard deviations (i.e. 69 percent confidence areas).

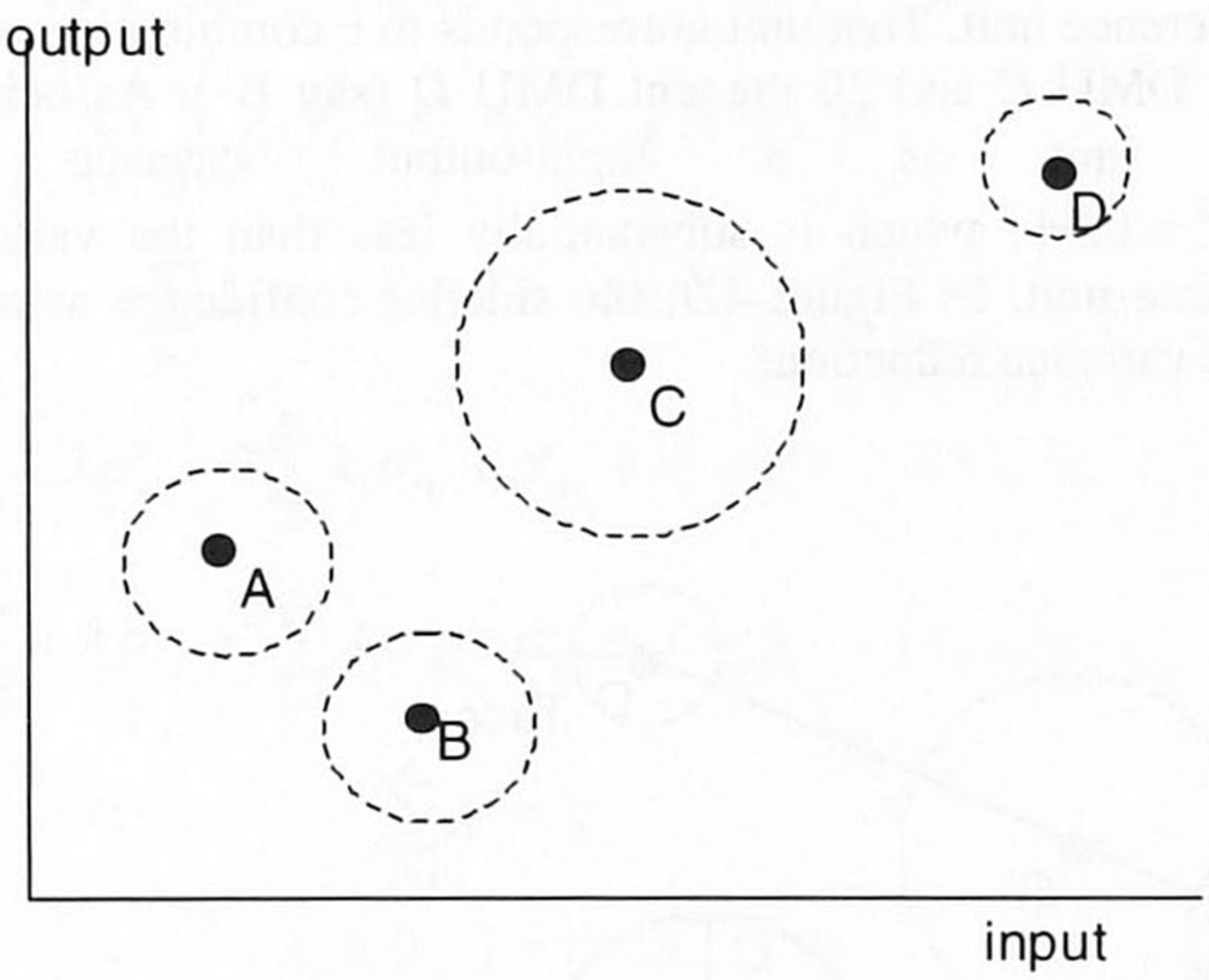


Figure 4. 7 Example data set and confidence areas

Figure 4.8 display the DEA frontier. The selection of an efficient reference unit for DMU *B* involves multiple optimal solutions. All solutions are particular convex combinations of DMUs *A*, *C* and *D*, and use 2 input and produce 3 output. However, not all solutions are equally risky. Suppose DEA selects an equaly weighed average of DMU *A* and DMU *C* (say *B'*) as a reference unit. Associated with this reference unit is an input-output variance of $0.5^2 \cdot 1^2 + 0.5^2 \cdot 2^2 = 1.25$. The diameter of the confidence area for reference unit *B'* reflects the maginitude of this variance.

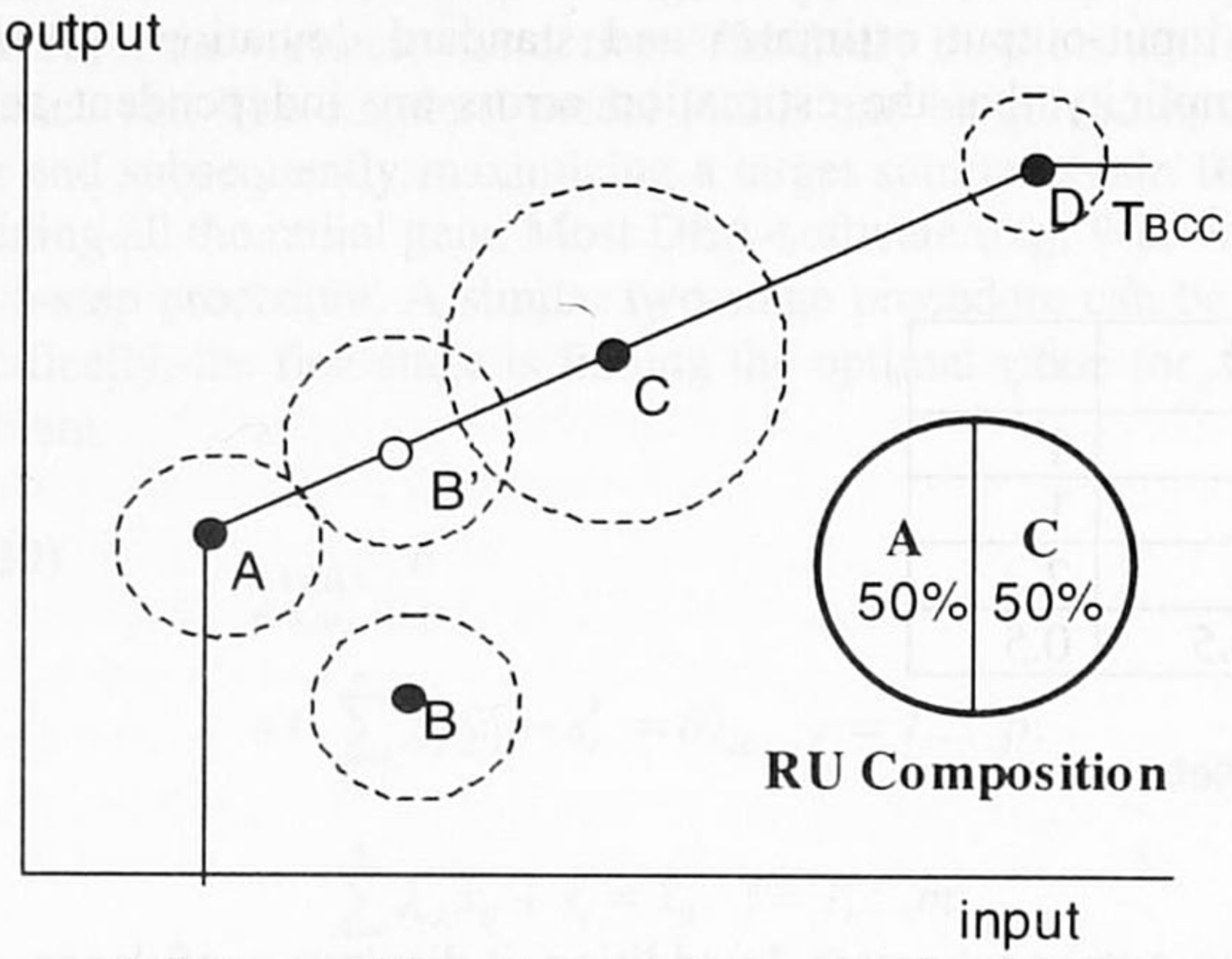


Figure 4. 8 A possible solution of the standard DEA model

Alternative reference units exist with less risk. The MVDEA approach is aimed at finding the minimum risk reference unit. That unit corresponds to a combination of 70 percent DMU *A*, 10 percent DMU *C* and 20 percent DMU *D* (say B"). Associated with that reference unit is a input-output variance of $0.7^2 \cdot 1^2 + 0.1^2 \cdot 2^2 + 0.2^2 \cdot 0.5^2 = 0.54$, which is substantially less than the variance found by the original referecne unit. In Figure 4.9, the smaller confidence area for reference unit *B"* reflects this variance reduction.

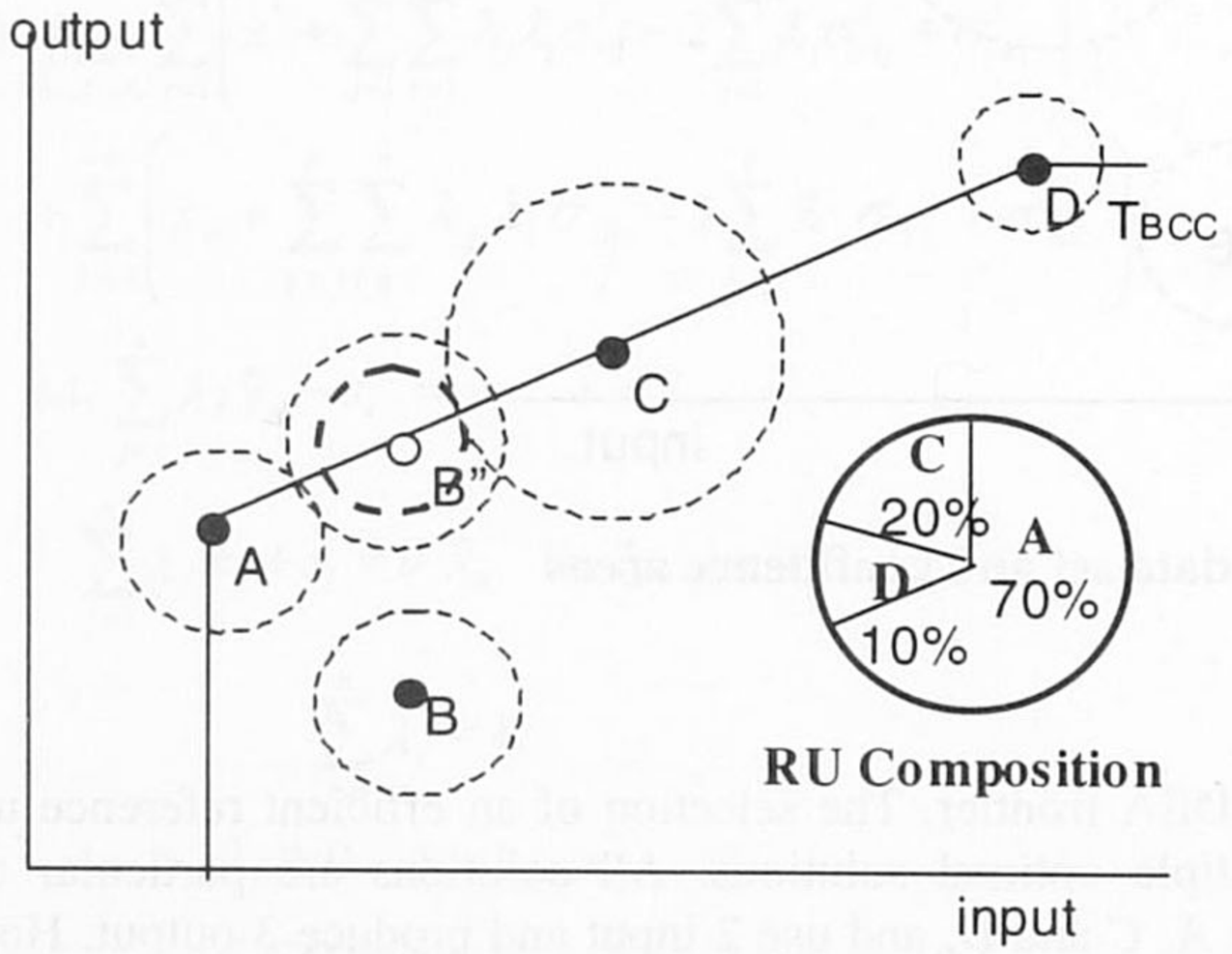


Figure 4. 9 The minimum variance solution

If the data set does not contain multiple reference units with the optimal input-output values, the variance correction does not have an effect on the optimal solution. However, if the data set does contain a multitude of such reference units, the selection of the one with least variance is prefererd by all risk-averse evaluators. Section 4 presents an empirical application for a real-life data set, for which the variance correction is important and substantially reduces error variance.

Restricting the reference units to the mean-variance set doesn't affect the efficiency estimate. The mean-variance set contains units with the same input-output estimates as the BCC refernce unit, but doesn't contain units that dominate that reference unit in terms of input-output estimates. Therefore, the maximal output augmentation is identical for both models. This implies that any bias of the BCC model is also present in a variance-corrected model. An alternative reference unit is selected that achieves the same output augemntation and consumes at most DMU_k's inputs, but with less uncertainty for its input-output vector relative to the evaluated unit. Restricting the error variance for the estimated input-output differences can further reduce that uncertainty. For example, the following model extends the standard model by requiring the output variance terms to be smaller or equal than γ'_r and the input variances to be smaller or equal than γ_i:

$$\textbf{(4. 22)} \quad \max_{\theta,\lambda_j,s^-,s'^-,d,d'} \theta + \varepsilon\left[\sum_{i=1}^{m} s_i + \sum_{r=1}^{p} s'_r + \sum_{i=1}^{m} d_i + \sum_{r=1}^{p} d'_r\right]$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j \hat{y}_{rj} - s'_r = \hat{y}_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i = \theta \hat{x}_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n}\sum_{l=1}^{n} \lambda_j \lambda_l \sigma'_{rlj} - 2\sum_{j=1}^{n} \lambda_j \sigma'_{rkj} + \sigma'_{rkk} + d'_r = \gamma'_r \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n}\sum_{l=1}^{n} \lambda_j \lambda_l \sigma_{ilj} - 2\sum_{j=1}^{n} \lambda_j \sigma_{ikj} + \sigma_{ikk} + d_i = \gamma_i \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$

$$s'_r, d'_r \geq 0 \quad r = 1,\ldots,p;$$

$$s_i, d_i \geq 0 \quad i = 1,\ldots,m.$$

In this expression, $d = (d_1 \cdots d_m)^T$ and $d' = (d'_1 \cdots d'_p)^T$ represent variance slacks.

However, the efficiency estimate may be affected by such restrictions, since the data set may not contain units that achieve the optimal output augmentation and consume at most DMUk's inputs and moreover have error variances that satisfy the restrictions. Since adding restrictions to a maximisation problem does imply a non-increased problem solution, the efficiency estimates would be lowered. If efficiency is underestimated, as is true in deterministic environments (see Chapter 1), this implies an increased estimation error. However, in stochastic environments the efficiency of the evaluated unit relative to the reference unit can be overestimated. We think such

overestimation occurs frequently in stochastic environments, because the DEA methodology is based on selecting extreme units. The overestimation of efficiency relative to the reference unit may cancel the effect of inefficient production by the reference units, and result in overestimation of efficiency relative to the production possibilities. In fact, a single outlier may cause substantial underestimation for all observations. If efficiency is overestimated, variance restrictions could reduce the estimation error.

Unfortunately, little is known about the distribution of DEA efficiency scores in stochastic environments. Therefore, we are not able to formally derive conditions under which variance restrictions improve efficiency estimates. However, we are currently analysing the implications of variance restrictions for efficiency estimation using Monte-Carlo simulations. Preliminary results suggest that variance restrictions (if properly specified) can substantially reduce bias and variance, especially if a large sample is available, if a large proportion of the observations operate efficiently and if the data set is heavily perturbed by stochastic disturbances. However, more research is needed to analyse the impact of variance restrictions on the statistical performance of DEA estimators.

Example 4.2 To illustrate the effects of variance restictions, consider 4 hypothetical DMUs with input-output estimates and standard deviations given in Table 4.2. Assume for simplicity that the estimation errors are independent across variables and across DMUs.

	$\hat{x}$	$\hat{y}$	σ_x	σ_y
A	1	2	1	1
B	2	1	1	1
C	3	3	2	2
D	5	3	0.5	0.5

Table 4. 2 Example data set

Figure 4.10 displays the input-output estimates and the associated confidence areas.

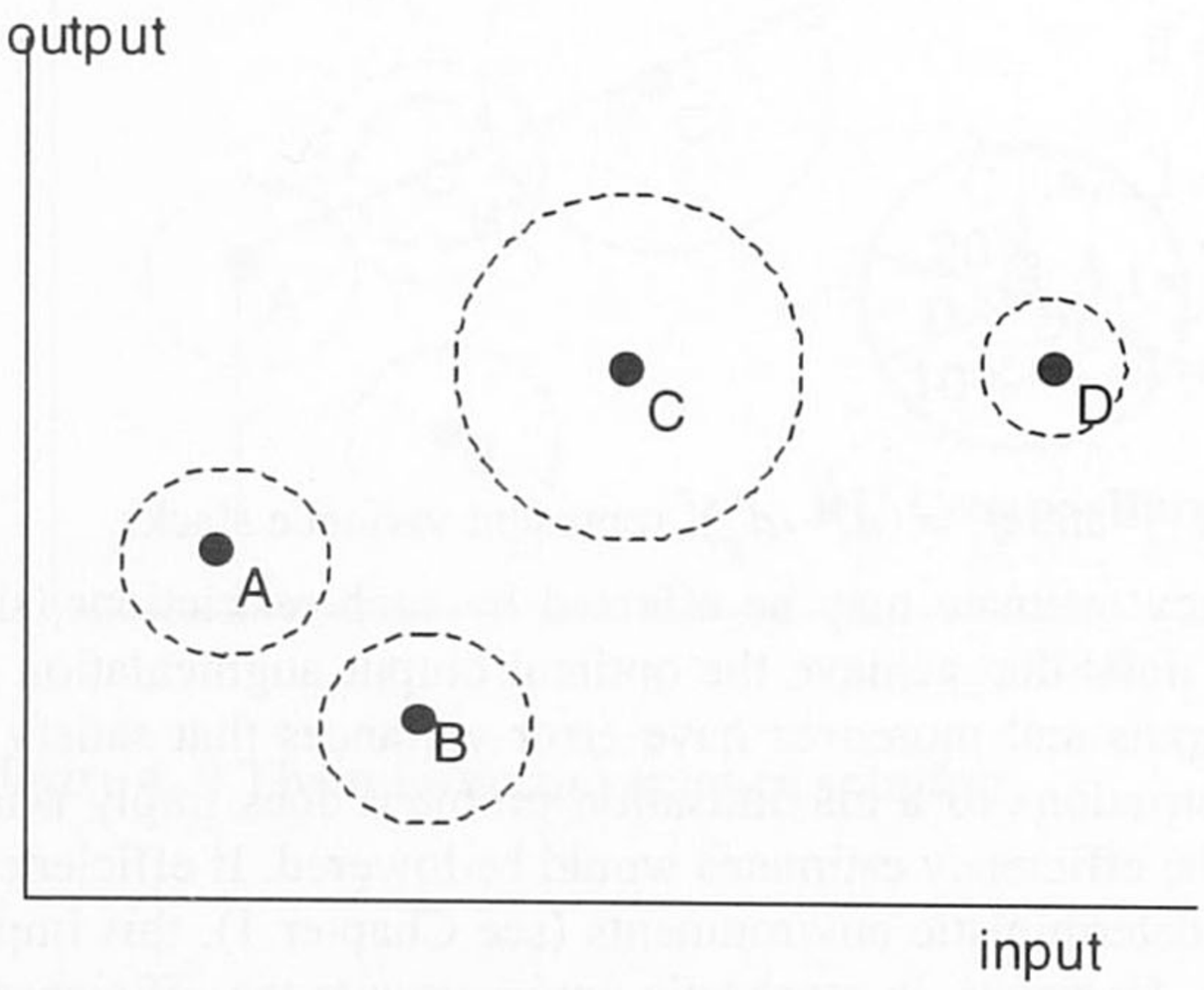


Figure 4. 10 Example data set and confidence areas

Figure 4.11 displays the DEA frontier. The selection of a efficient reference unit for DMU *B* involves just a single solution. That solution is an equally weighed average of DMU *A* and DMU *C* (say *B'*). That reference unit uses 2 input and produces 3 output. Associated with this reference unit is a input-output varance of
$0.5^2 \cdot 1^2 + 0.5^2 \cdot 2^2 = 1.25$.

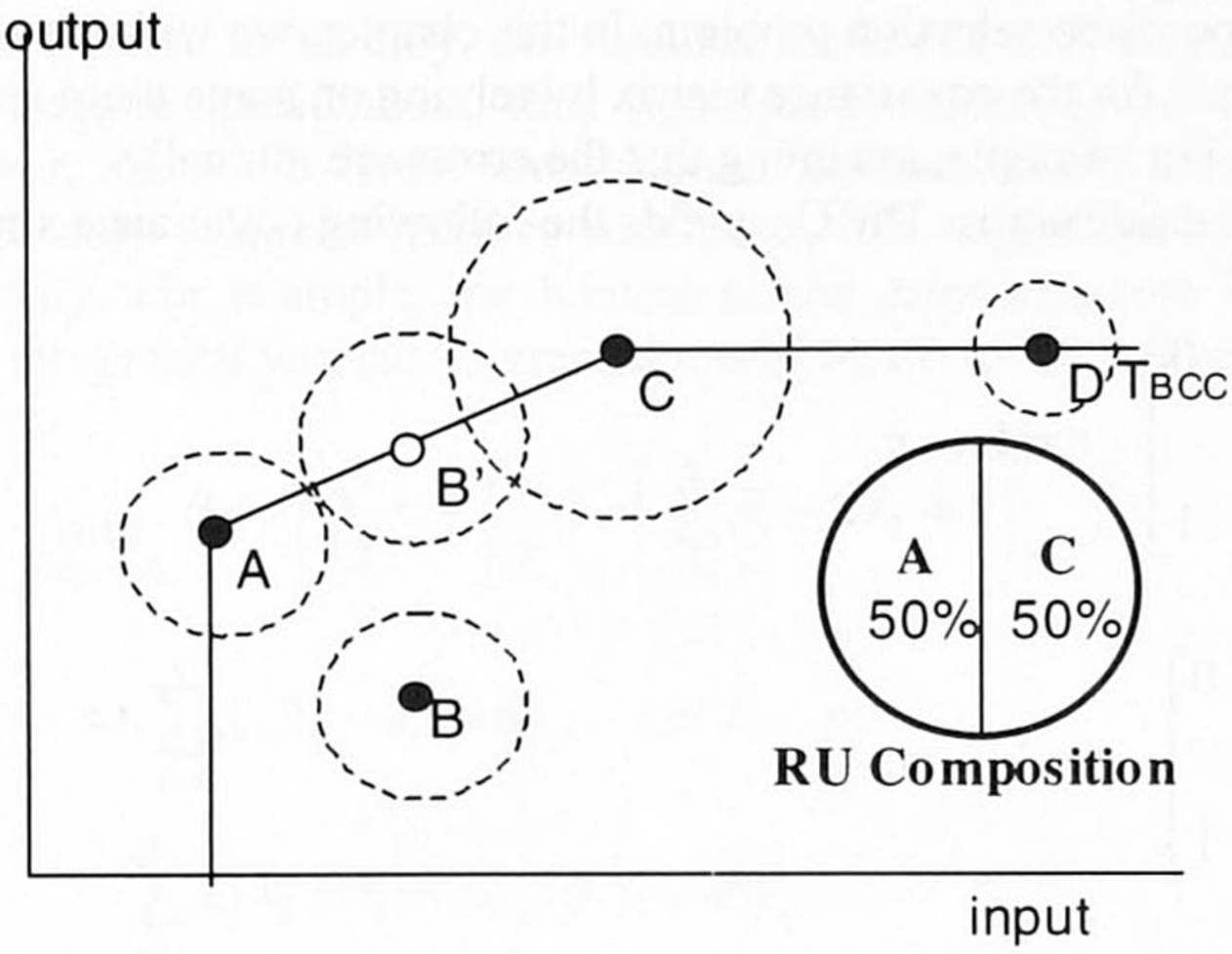


Figure 4. 11 The standard DEA solution

Imposing variance restrictions can reduce the variance. For example, restricting the input-output variance not to exceed 0.56 gives a combination of 66 percent DMU *A*, 17 percent DMU *C* and 27 percent DMU *D* (say *B''*). Associated with that reference unit that consumes 2 input and produces 2.33 output. Figure 4.12 displays the associated variance reduction and the associated change in input-output vector.

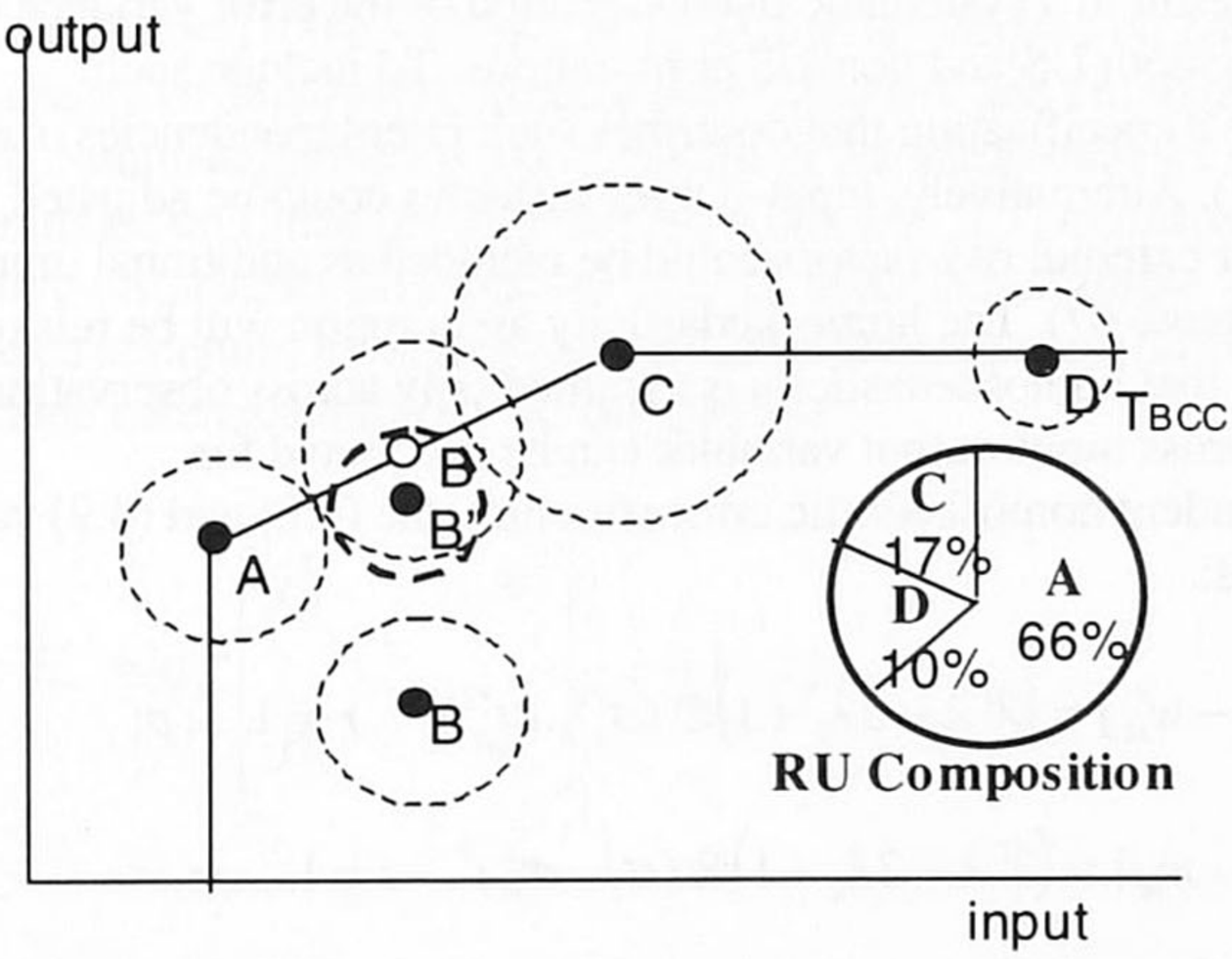


Figure 4. 12 The solution of a variance restricted model

4.4. Special cases

The above framework for including parameter uncertainty requires the covariance matrix of the estimation errors to be specified accurately. The specification could be based on prior theory or empirical estimation. However, in many research environments, neither theory nor data allow for an accurate specification. The resulting uncertainty for the covariance matrix could be incorporated into the analysis analogously to the uncertainty for the input-output variables. For example, Bawa and Klein (1976, 1977) incorporated the estimation risk of using the sample covariance matrix as an estimate for the true covariance matrix in the portfolio selection problem. In this chapter, we will circumvent the information requirement for the covariance matrix by relying on some more specific distribution assumptions. For example, assuming that the errors are mutually independent and homoskedastic across DMUs, yields the following covariance structure:

$$\textbf{(4. 23)}\quad \Sigma'_r = \sigma'^2_r \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \quad r = 1,\ldots, p$$

$$\textbf{(4. 24)}\quad \Sigma_i = \sigma_i^2 \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \quad r = 1,\ldots, m.$$

In this expression, σ'^2_r (σ_i^2) represents a scalar variance term that applies for all DMUs, i.e. $\sigma'_{r11} = \ldots = \sigma'_{rnn} = \sigma'^2_r$ ($\sigma_{i11} = \ldots = \sigma_{inn} = \sigma_i^2$).

The independence assumption excludes effects from correlation between DMUs. In our opinion, this assumption is tenable for idiosyncratic measurement error, random noise and outliers. Nevertheless, interdependencies can occur if multiple DMUs are influenced by a common external source of risk. For example, an unanticipated appreciation of the US dollar generally simultaneously lowers revenues for firms that export from the US and raises revenues for firms that import from the US. This causes a negative correlation between revenues of US and non-US corporations. Ignoring this interdependency would result in a systematic overestimation of the error variance of reference units containing both US and non-US corporations. To include such correlation would require a specification that describes such interdependencies, such as in the general model (4.7). Alternatively, input-output variables could be adjusted for systematic uncertainty, or external risk factors could be included as additional input-output variables (see footnote 67). The homoskedasticity assumption will be relaxed hereafter. Note however, that homoskedasticity is assumed only across observations, and that heteroskedasticity across input-output variables can be accounted for.

Using the independent homoskedastic error structure, the (4.8) and (4.9) variance terms can be expressed as:

$$\textbf{(4. 25)}\quad \sigma'^2_{\cdot k}(\lambda) = V[w'_{r\cdot}\lambda - w'_{rk}] = \left(\lambda^T\lambda - 2\lambda_k + 1\right) \otimes (\sigma'^2_1 \ldots \sigma'^2_p)^T \quad r = 1,\ldots, p;$$

$$\textbf{(4. 26)}\quad \sigma^2_{\cdot k}(\lambda) = V[w_{i\cdot}\lambda - w_{ik}] = \left(\lambda^T\lambda - 2\lambda_k + 1\right) \otimes (\sigma^2_1 \ldots \sigma^2_m)^T \quad i = 1,\ldots, m.$$

Consequently, the mean-variance admissible set amounts to:

$$\textbf{(4. 27)}\ A_{MV} = \left\{ \lambda \in \Re_+^n \middle| \left\{ \tau \in \Re_+^n \middle| \begin{pmatrix} \hat{Y}\tau \\ -\hat{X}\tau \end{pmatrix} - \begin{pmatrix} s' \\ s \end{pmatrix} = \begin{pmatrix} \hat{Y}\lambda \\ -\hat{X}\lambda \end{pmatrix}; \tau^T\tau - 2\tau_k + d = \lambda^T\lambda - 2\lambda_k; \right.\right.$$
$$\left.\left. \tau^T e = 1; s', d', s, d \geq 0; e'^T s' + e^T s + d > 0 \right\} = \varnothing; \lambda^T e = 1 \right\}.$$

Since the structure of the mean-variance admissible set depends on the error variance terms in relative terms only, the variance terms of (4.15) cancel out, and consequently, do not require specification, thus significantly reducing the information requirement. Moreover, since the errors are assumed homoskedastic across observations this set involves only a single variance restriction, thus significantly reducing computational complexity. For example, the homoskedastic error structure in (4.23) and (4.24) can reduce the general variance-corrected model (4.14) to the following model:

$$\textbf{(4. 28)} \quad \min_{\theta,\lambda_j,s_i,s'_r} \theta + \varepsilon\left[\sum_{i=1}^{m} s_i + \sum_{r=1}^{p} s'_r - \left(\sum_{j=1}^{n} \lambda_j^2 - 2\lambda_k + 1\right)\right]$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j \hat{y}_{rj} - s'_r = \theta \hat{y}_{rk} \quad r = 1,\ldots,p;$$

$$\sum_{j=1}^{n} \lambda_j \hat{x}_{ij} + s_i = \hat{x}_{ik} \quad i = 1,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,n;$$
$$s'_r \geq 0 \quad r = 1,\ldots,p;$$
$$s_i \geq 0 \quad i = 1,\ldots,m.$$

A potential limitation of this approach is that it may discriminate against small-sized DMUs. The covariance structure in (4.23) and (4.24) assumes homoskedasticity across DMUs. Therefore, if the variance terms positively depend on DMU size, the relative error variance of small-sized units is overestimated, and consequently large mean-variance dominated units may be included in the admissible set and small mean-variance nondominated unit may be excluded from that set. To circumvent this problem, alternative disturbance distributions may be used. For example, employing the following covariance structure, involving standard deviations that are proportional to the estimates can include heteroskedasticity across DMUs:

$$\textbf{(4. 29)}\ \Sigma'_r = \sigma'^2_r \begin{bmatrix} \hat{y}_{r1}^2 & & 0 \\ & \ddots & \\ 0 & & \hat{y}_{rm}^2 \end{bmatrix} \quad r = 1,\ldots,p;$$

$$\textbf{(4. 30)}\ \ \Sigma_i = \sigma_i^2 \begin{bmatrix} \hat{x}_{i1}^2 & & 0 \\ & \ddots & \\ 0 & & \hat{x}_{in}^2 \end{bmatrix} \quad i = 1,\ldots,m.$$

Using this covariance structure, the (4.8) and (4.9) variance terms can be expressed as:

$$\textbf{(4. 31)}\ \ \sigma'^2_{\cdot k}(\lambda) = \begin{pmatrix} \sigma'^2_1 \\ \vdots \\ \sigma'^2_p \end{pmatrix} \otimes \left[\begin{pmatrix} \lambda^T(\hat{y}_{1\cdot}\hat{y}_{1\cdot}^T)\lambda \\ \vdots \\ \lambda^T(\hat{y}_{p\cdot}\hat{y}_{p\cdot}^T)\lambda \end{pmatrix} - (2\lambda_k - 1) \begin{pmatrix} \hat{y}_{1k}^2 \\ \vdots \\ \hat{y}_{pk}^2 \end{pmatrix} \right] \quad r = 1,\ldots,p;$$

$$\textbf{(4. 32)}\ \ \sigma^2_{\cdot k}(\lambda) = \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_m^2 \end{pmatrix} \otimes \left[\begin{pmatrix} \lambda^T(\hat{x}_{1\cdot}\hat{x}_{1\cdot}^T)\lambda \\ \vdots \\ \lambda^T(\hat{x}_{m\cdot}\hat{x}_{m\cdot}^T)\lambda \end{pmatrix} - (2\lambda_k - 1) \begin{pmatrix} \hat{x}_{1k}^2 \\ \vdots \\ \hat{x}_{mk}^2 \end{pmatrix} \right] \quad i = 1,\ldots,m.$$

Consequently, the mean-variance admissible set amounts to:

$$\textbf{(4. 33)}\ \ A_{MV} = \left\{ \lambda \in \Re_+^n \middle| \left\{ \tau \in \Re_+^n \middle| \begin{pmatrix} \hat{Y}\tau \\ -\hat{X}\tau \end{pmatrix} - \begin{pmatrix} s' \\ s \end{pmatrix} = \begin{pmatrix} \hat{Y}\lambda \\ -\hat{X}\lambda \end{pmatrix}; \right.\right.$$

$$\begin{pmatrix} \tau^T(\hat{y}_{1\cdot}\hat{y}_{1\cdot}^T)\tau \\ \vdots \\ \tau^T(\hat{y}_{p\cdot}\hat{y}_{p\cdot}^T)\tau \end{pmatrix} - 2\tau_k \begin{pmatrix} \hat{y}_{1k}^2 \\ \vdots \\ \hat{y}_{pk}^2 \end{pmatrix} + d' = \begin{pmatrix} \lambda^T(\hat{y}_{1\cdot}\hat{y}_{1\cdot}^T)\lambda \\ \vdots \\ \lambda^T(\hat{y}_{p\cdot}\hat{y}_{p\cdot}^T)\lambda \end{pmatrix} - 2\lambda_k \begin{pmatrix} \hat{y}_{1k}^2 \\ \vdots \\ \hat{y}_{pk}^2 \end{pmatrix};$$

$$\begin{pmatrix} \tau^T(\hat{x}_{1\cdot}\hat{x}_{1\cdot}^T)\tau \\ \vdots \\ \tau^T(\hat{x}_{m\cdot}\hat{x}_{m\cdot}^T)\tau \end{pmatrix} - 2\tau_k \begin{pmatrix} \hat{x}_{1k}^2 \\ \vdots \\ \hat{x}_{mk}^2 \end{pmatrix} + d = \begin{pmatrix} \lambda^T(\hat{x}_{1\cdot}\hat{x}_{1\cdot}^T)\lambda \\ \vdots \\ \lambda^T(\hat{x}_{m\cdot}\hat{x}_{m\cdot}^T)\lambda \end{pmatrix} - 2\lambda_k \begin{pmatrix} \hat{x}_{1k}^2 \\ \vdots \\ \hat{x}_{mk}^2 \end{pmatrix};$$

$$\left.\left. \tau^T e = 1; s', d', s, d \geq 0; e'^T(s' + d') + e^T(s + d) > 0 \right\} = \varnothing; \lambda^T e = 1 \right\}.$$

As in the homoskedastic model, the individual variance terms cancel out, and consequently do not require specification. However, the heteroskedastic error structure involves *m+s* variance conditions, and therefore involves more computational burden than the homoskedastic structure.

$$\textbf{(4. 34)} \qquad \max_{\theta,\lambda_j,s_i,s'_r} \theta + \varepsilon\left[\sum_{i=1}^{m} s_i + \sum_{r=1}^{p} s'_r\right]$$

$$+\varepsilon\left[\sum_{j=1}^{n}\lambda_j^2\left(\sum_{i=1}^{m}\hat{x}_{ij}^2+\sum_{r=1}^{p}\hat{y}_{rj}^2\right)-(2\lambda_k-1)\left(\sum_{i=1}^{m}\hat{x}_{ik}^2+\sum_{r=1}^{p}\hat{y}_{rk}^2\right)\right]$$

$$\text{s.t.} \sum_{j=1}^{n}\lambda_j \hat{y}_{rj} - s'_r = \theta\hat{y}_{rk} \quad r=1,\ldots,p;$$

$$\sum_{j=1}^{n}\lambda_j \hat{x}_{ij} + s_i = \hat{x}_{ik} \quad i=1,\ldots,m;$$

$$\sum_{j=1}^{n}\lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j=1,\ldots,n;$$

$$s'_r \geq 0 \quad r=1,\ldots,p;$$

$$s_i \geq 0 \quad i=1,\ldots,m.$$

4.5. Empirical application

This section applies the mean-variance approach to a real-life data set of large Indonesian commercial banks. Performance evaluation of financial institutions is a research area in which DEA has seen considerable application. For example, in a survey of frontier efficiency studies for financial institutions, Berger and Humphrey (1997) find that 69 out of 122 studies use DEA. Unfortunately, there is no generally accepted model for the technology of financial institutions. In general, the technology is represented by certain outputs (e.g. loans, mortgages and insurance), produced by certain inputs (e.g. personnel, workspace and computers), in a microenvironment that is characterised by a certain number of environmental variables (e.g. number of competitors, consumer demographics, and the prices for the inputs and outputs). However, the choice of the variables and the assumptions imposed on the relationships between the variables differ greatly.

This study follows Yue (1992) and Yeh (1996) in specifying the relevant input-output variables[70]. The output variables used are:

(1) *Total Net Loans* (TNL)[71],

(2) *Interest Income* (II), and

(3) *Other Income* (OI)[72].

---

[70] For a survey of recent research on performance of financial institutions, see Berger and Humphrey (1997).

[71] Defined as: Loans and advances plus secured loans minus loan loss reserves.

[72] Defined as: Commission fees plus other operating income.

The input variables used are:

(1) *Customer Deposits* (CD)[73],

(2) *Interest Expenses* (IE), and

(3) *Other Expenses* (OE)[74].

That data set contains data on these variables for 49 banks, taken from the income statements over the year 1995 and the balance sheet of the end of year 1995. Some descriptive statistics of the data set are displayed in Table 4.3.

	Total Net Loans	Interest Income	Other Income	Customer Deposits	Interest Expenses	Other expenses
Mean	3222.57	556.59	44.89	3384.73	406.90	142.93
Median	861.30	183.40	11.40	884.00	127.60	40.30
Maximum	21826.60	3990.00	447.00	22562.80	2562.00	1494.40
Minimum	148.00	48.70	1.30	161.90	35.50	7.40
Std. Dev.	5432.77	871.19	82.85	5392.75	633.83	271.45
Skewness	2.28	2.32	3.04	2.11	2.18	3.32
Kurtosis	7.14	7.81	13.23	6.48	6.77	14.97

Table 4. 3 Descriptive statistics of the data set

Starting point for the analysis was the BCC model (1.10). To account for parameter uncertainty, we also used a variance-corrected model. As discussed above, such a model circumvents the selection of stochastically dominated reference units.

Since no additional information was available, the input-output variables were estimated using the observed values. In addition, we assumed that the estimation errors are mutually independent. As discussed in the previous section, the independence assumption is plausible for measurement error, random noise and outliers, and moreover substantially reduces the information requirement and the complexity of the admissible set. Moreover, we assumed the heteroskedastic structure (4.23) and (4.24), giving model (4.28). This was because, as in most reported bank studies, the Indonesian banking industry is characterised by a small number of big institutions and a large number of mid-sized and small-sized institutions. The descriptive statistics illustrate the wide variation and the skewness of the size distribution. The complete specification for the mean-variance model is as follows:

---

[73] Defined as: Demand deposits plus time deposits plus saving deposits

[74] Defined as: Payroll expenses plus loan loss provisions plus other operating expenses

$$\textbf{(4. 35)} \quad \underset{\theta,\lambda,s'_{TNL},s'_{II},s'_{OI},s_{CD},s_{IE},s_{OE}}{\text{Max}} \quad \theta + \varepsilon\left[s'_{TNL} + s'_{II} + s'_{OI} + s_{CD} + s_{IE} + s_{OE}\right]$$

$$- \varepsilon[\sum_{j=1}^{49} \lambda_j^2 (TNL_j^2 + II_j^2 + OI_j^2 + CD_j^2 + IE_j^2 + OE_j^2)$$

$$- (\lambda_k - 1)(TNL_k^2 + II_k^2 + OI_k^2 + CD_k^2 + IE_k^2 + OE_k)]$$

$$\text{s.t.} \sum_{j=1}^{49} \lambda_j TNL_j - s'_{TNL} = \theta TNL_k;$$

$$\sum_{j=1}^{49} \lambda_j II_j - s'_{II} = \theta II_k;$$

$$\sum_{j=1}^{49} \lambda_j OI_j - s'_{OI} = \theta OI_k;$$

$$\sum_{j=1}^{49} \lambda_j CD_j + s_{CD} = CD_k;$$

$$\sum_{j=1}^{49} \lambda_j IE_j + s_{IE} = IE_k;$$

$$\sum_{j=1}^{49} \lambda_j OE_j + s_{OE} = OE_k;$$

$$\sum_{j=1}^{49} \lambda_j = 1;$$

$$\lambda_j \geq 0 \quad j = 1,\ldots,49;$$

$$s'_{TNL}, s'_{II}, s'_{OI}, s_{CD}, s_{IE}, s_{OE} \geq 0.$$

Both this mean-variance model and its traditional counterpart (i.e. model (4.35) excluding the variance correction in the objective function) were solved.

Table 4.4 displays the efficiency estimates for all 49 banks, i.e the optimal values for θ. The estimates refer to both the BCC model and the variance-corrected model, because the variance correction does not affect the efficiency estimate. This is because, as discused above, for each unit included in the BCC set, but excluded in the mean-variance set, the latter contains an alternative reference unit with identical input-output estimates, but less parameter uncertainty.

However, if multiple optimal solutions exist, the variance correction can reduce parameter uncertainty. The introduction of the variance correction into the model specification led to variance reduction relative to the traditional model for 17 banks.

Bank	Efficiency	Bank	Efficiency
1	1.000	26	1.147
2	1.000	27	1.153
3	1.000	28	1.255
4	1.015	29	1.322
5	1.000	30	1.174
6	1.263	31	1.000
7	1.000	32	1.185
8	1.000	33	1.599
9	1.000	34	1.140
10	1.000	35	1.302
11	1.000	36	1.026
12	1.000	37	1.429
13	1.000	38	1.230
14	1.068	39	1.000
15	1.000	40	1.081
16	1.000	41	1.144
17	1.173	42	1.032
18	2.100	43	1.297
19	1.000	44	1.456
20	1.000	45	1.120
21	1.000	46	1.212
22	1.000	47	1.269
23	1.000	48	1.000
24	1.000	49	1.315
25	1.000		

Table 4. 4 Standard DEA efficiency scores

For these banks, a selection of the DEA output is displayed in Table 4.5. First, the Table displays the variance correction factor that is included in the objective of the mean-variance model, i.e.:

$$\textbf{(4. 36)}\quad \sum_{j=1}^{49} \lambda_j^2 (TNL_j^2 + II_j^2 + OI_j^2 + CD_j^2 + IE_j^2 + OE_j^2) - (\lambda_k - 1)(TNL_k^2 + II_k^2 + OI_k^2 + CD_k^2 + IE_k^2 + OE_k)].$$

This correction factor excludes the (4.23) and (4.24) variance scalars $\sigma_r'^2$, r=1,…,p and σ_i^2, i=1,…,m. Since utility is assumed monotonously decreasing in error variance, these terms do not require specification to exclude stochastically dominated reference units. However, not specifying the variance scalars implies that the correction factor doesn't represent the actual total variance for the referencing inputs and outputs. The variance reduction may be overestimated (underestimated) if the input-output variables for which the highest reduction is achieved have the lowest (highest) error variance. The traditional model selects reference units that are dominated

in mean-variance terms. The data set contains other units that use the same minimal input fraction to produce at least the outputs of the evaluated unit, but with less uncertainty. Second, the table displays the number of referencing units.

	BCC		MV-DEA		
Bank	**Variance Correction Factor**	**Number of RUs**	**Variance Correction Factor**	**Number of RUs**	**Error Variance Reduction**
6	111043	2	55846	5	49.71%
17	84064	4	79006	5	6.02%
18	409770	3	122134	6	70.19%
27	330747	5	324436	6	1.90%
28	77063	5	72721	7	5.63%
29	2345	4	2344	5	0.01%
30	991655	5	934003	6	5.81%
33	144228	5	89419	7	38.00%
35	113908	4	101074	6	11.27%
36	526616	5	522761	6	0.73%
37	18484989	4	18421539	5	0.34%
38	259019	5	192946	8	25.51%
40	82834	4	70997	6	14.29%
41	325031	5	321296	6	1.15%
43	464299	5	336235	8	27.58%
44	2765578	6	2695803	7	2.52%
48	950757	6	948849	7	0.20%

Table 4. 5 Effects of the variance correction

For Bank 18, the composition of the reference unit and the estimated values and error variances for the referencing input-output variables are displayed in Table 4.5. The traditional model selected a convex combination of bank 15, bank 20 and bank 31 as a reference unit. In this reference unit, bank 31 represents more than 90 percent. Therefore, the reference unit is poorly diversified and consequently involves relatively much error variance. By contrast, the mean-variance model selects a composite of Bank 15, Bank 16, Bank 20, Bank 31, Bank 39 and Bank 48 as a reference unit. That reference unit is better diversified, and consequently less uncertain. More specifically, that unit involves only 26 to 53 percent of the error variance associated with the traditional reference unit. A follow-up inspection of the data revealed that a particular convex combination of Bank 16, Bank 39, and Bank 48 gives the same input-output combinations as Bank 31. This means that additional empirical support can be found for the attainability of the input-output estimates for the BCC reference unit. The mean-variance reference unit not only dominates Bank4 in a deterministic setting, but also stochastically dominates by

second order, because it has both more favourable input-output estimates and less error variance.

	BCC			MV-DEA		
	Ref. Unit Composition	**Estimated value**	**Error Variance**	**Ref. Unit Composition**	**Estimated value**	**Error Variance**
TNL	0.024 Bank 15	735.15	364016.23	0.024 Bank 15	735.15	100223.65
II	0.074 Bank 20	126.20	8790.03	0.203 Bank 16	126.20	2719.37
OI	0.902 Bank 31	74.70	2520.43	0.074 Bank 20	74.70	963.80
CD		13.28	129.83	0.187 Bank 31	13.28	34.96
IE		18.19	169.21	0.245 Bank 39	18.19	59.03
OE		295.55	34146.84	0.276 Bank 48	295.55	18137.86

Table 4. 6 Effects of the variance correction for Bank 18

4.6. Monte-Carlo simulations

As discussed in Section 4.3, including variance *corrections* can reduce uncertainty about the performance of the evaluated DMU relative to the reference unit. Variance *restrictions* can further reduce that uncertainty. However, in contrast to variance corrections, variance restrictions can interfere with the efficiency estimates as well. In fact, one important motivation for using variance restrictions is the expectation that well-diversified reference units can give better efficiency estimates than ill-diversified reference units. However, variance restrictions do not necessarily improve efficiency estimation. Variance restrictions restrict the uncertainty of performance relative to the reference unit. It can be impossible to find efficient reference units that satisfy the restrictions, if the data set contains few efficient DMUs. It is interesting to analyse the resulting trade-off. Unfortunately, no analytical solutions are available for the statistical distribution of DEA estimates in stochastic environments. Consequently, conditions under which variance restrictions improve efficiency estimates cannot be derived analytically. Therefore, we analysed some aspect of the statistical distribution using Monte-Carlo simulations.

As discussed in section 3.7, there currently is no generally accepted framework for performing simulation studies in DEA. Therefore, we adopt a heuristic approach in this study. Specifically, we use a data generating process that is relatively simple, and hence limits the computational burden, while allowing us to answer our research questions[75].

[75] In this study, we systematically vary the sample size, the noise level, and the inefficieny level. However, for simplicity, we focus on a specific production technology, a specific distribution for the observations and a specific variance restriction. We are currently working on a simulation study that includes the effects of technology and distribution. Prelilimary results suggest that the above model gives a fair representation of the potential of MV-DEA. In addition, that study considers the specification of the best variance restriction.

Specifically, in the simulations, all units were assumed to face a common production technology, which is represented by the following stochastic frontier model:

$$\textbf{(4. 37)}\quad y_j = \left| f(x_{1j}, x_{2j}) - u_j + w_j, \right|$$
$$f(x_{1j}, x_{2j}) = 100 + 0.4x_{1j} + 0.6x_{2j},$$
$$u_j \sim \left| N(0, \sigma_u) \right|,$$
$$w_j \sim N(0, \sigma_v),$$
$$j = 1, \ldots, n.$$

In this specification, $f(x_{1j}, x_{2j})$ represents the maximum output associated with the inputs of the j-th unit, x_{1j} and x_{2j}. The inputs follow a bivariate lognormal distribution with (for both inputs) mean 100 and standard deviation 10, and coefficient of correlation 0.9. The lognormal distribution was selected to prevent negative values and moreover to reflect the skewness of the size distribution typically encountered in empirical studies. Similarly, the correlation was introduced so as to reflect the high degree of correlation between inputs.

Observed output y_j differs from the maximal output by inefficiency u_j and stochastic disturbance w_j. The efficiencies follow an independent and identical half normal distribution; i.e. they are the absolute values of a variable with mean 0 and standard deviation σ_u[76]. Noise is independently and identically normal distributed with mean 0 and standard deviation σw.

This study compares the BCC model (1.20) with a mean-variance model that includes a variance correction in the objective. In addition, the mean-variance model includes a restriction on the variance of the estimation errors for the cost differences. Since no additional information is available, the observations were used to estimate the unperturbed outputs, i.e. $y_j = f(x_{1j}, x_{2j}) - u_j$. Consequently, the estimation errors coincide with the disturbances, and the error variance amounts to:

$$\textbf{(4. 38)}\quad (\sum_{j=1}^{n} \lambda_j^2 - 2\lambda_k + 1)\sigma_w^2.$$

We restricted this term to be smaller than 10 percent of the variance of the disturbances σ_w^2. By restricting variance in relative terms, we circumvent the specification of the noise level σ_w^2. Specifically, the following variance restriction was included:

$$\textbf{(4. 39)}\quad (\sum_{j=1}^{n} \lambda_j^2 - 2\lambda_k + 1) \leq 0.1.$$

The complete specification for the mean-variance model amounts to:

---

[76] Under this assumption, ineffciency has an average value of $(2/\pi)^{0.5}\sigma_u$ and a standard deviation of $((\pi-2)/\pi)^{0.5}\sigma_u$.

$$\textbf{(4. 40)} \quad \max_{\theta,\lambda_j,s_1',s_1,s_2,d_1'} \theta + \varepsilon(s_1' + s_1 + s_2 + d_1')$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_j - s_1' = \theta y_k;$$

$$\sum_{j=1}^{n} \lambda_j x_{1j} + s_1 = x_{1k};$$

$$\sum_{j=1}^{n} \lambda_j x_{2j} + s_2 = x_{2k};$$

$$\sum^{n} \lambda_j^2 - 2\lambda_k + 1 + d_1' = 0.1;$$

$$\sum_{j=1}^{n} \lambda_j = 1;$$

$$\lambda_j, s_1', s_1, s_2, d_1' \geq 0 \quad j = 1,\ldots,n.$$

Starting point for the simulations was a setting with 200 observations, and standard deviations for inefficiency and noise of 10, i.e. 5 percent of average output. To analyze the influence of the sample size and the noise and inefficiency levels, we subsequently varied the values for n, σ_u and σ_w. More specifically, we considered sample sizes of 50, 200 and 500, and values of 1,5 and 10 percent for both noise and inefficiency.

Each simulation consisted of generating a set of artificial data from the above data-generating process, and employing these data to create the efficiency estimators from both the BCC model and the mean-variance model. For each of the combinations of sample size, inefficiency and noise, we did 1000 simulations. Next, these estimates were used to gauge the sampling distribution properties of the estimators of the two competing models. We considered three sample statistics: bias (BIAS), standard deviation (STD) and root mean squared error (RMSE). These statistics were computed as follows:

$$\textbf{(4. 41)} \quad BIAS = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j);$$

$$\textbf{(4. 42)} \quad STD = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \bar{\hat{\theta}}_j)^2;$$

$$\textbf{(4. 43)} \quad RMSE = \tfrac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j)^2.$$

In these expressions, θ_j represents the true Debreu-Farrell output efficiency measure (1.3). For frontier model (4.37), the latter corresponds to:

$$\textbf{(4. 44)} \quad \theta_j = \frac{f(x_{1j}, x_{2j})}{y_j}.$$

For the above statistics, the mean values over the replications were computed. The results are summarised in Table 4.6. For the base case (200 observations, 5 percent inefficiency and 5 percent noise), the estimates are centred well above true efficiency with a bias of -0.066. In addition, the estimation errors are widely spread out with a standard deviation of 0.037. This implies that the BCC reference units generally suggest performance improvements that are not attainable under the production technology, and, moreover, that are highly uncertain. Imposing the variance restriction reduces the bias to 0.002 and the standard deviation to 0.019. Apparently, the data set generally contains well-diversified reference units that reflect the production possibilities better that the BCC reference units, which are less well diversified. The distribution of the estimation errors for the BCC model and the mean-variance model are displayed in Figure 4.13.

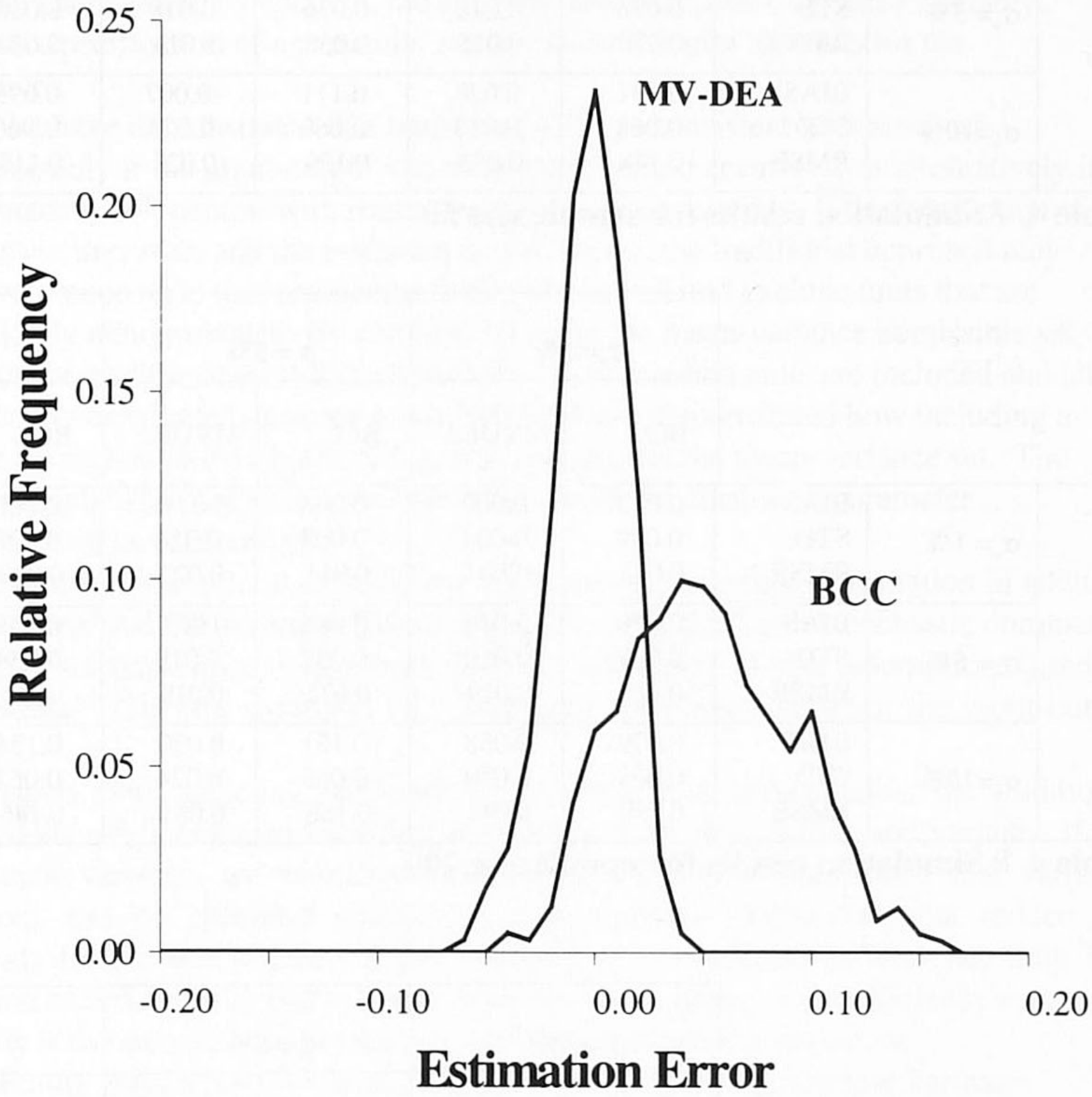


Figure 4. 13 Sample distribution of estimation errors $(\hat{\theta}_j - \theta_j)$ for the two competing models

Similar results are obtained for the scenario's involving other sample sizes and inefficiency and noise levels. For most scenario's the mean-variance model gives better estimates. However, if the data set is small, inefficiency is high and noise is low, the traditional approach outperforms the mean-variance model. By contrast, the larger the sample, the higher noise and the lower inefficiency, the more effective the variance

restriction. Apparently, if many observations are located near the true frontier, the data set contains well-diversified reference unit that gives a good representation of the production possibilities.

			σ_u= 1%		σ_u= 5%		σ_u= 10%	
			BCC	MVDEA	BCC	MVDEA	BCC	MVDEA
	σ_w= 1%	BIAS	0.008	-0.002	0.001	-0.020	-0.002	-0.035
		STD	0.008	0.004	0.009	0.015	0.010	0.024
		RMSE	0.011	0.004	0.009	0.025	0.010	0.042
	σ_w= 5%	BIAS	0.060	0.011	0.045	-0.008	0.035	-0.029
		STD	0.036	0.010	0.036	0.018	0.038	0.030
		RMSE	0.070	0.015	0.058	0.019	0.051	0.041
	σ_w=10%	BIAS	0.131	0.028	0.111	-0.007	0.095	-0.015
		STD	0.068	0.018	0.069	0.024	0.066	0.033
		RMSE	0.148	0.033	0.129	0.025	0.115	0.036

Table 4. 7 Simulation results for sample size 50

			σ_u= 1%		σ_u= 5%		σ_u= 10%	
			BCC	MVDEA	BCC	MVDEA	BCC	MVDEA
	σ_w= 1%	BIAS	0.013	0.000	0.008	-0.017	0.005	-0.035
		STD	0.009	0.004	0.008	0.014	0.008	0.025
		RMSE	0.015	0.004	0.011	0.022	0.010	0.043
	σ_w= 5%	BIAS	0.078	0.016	0.066	-0.002	0.055	-0.025
		STD	0.037	0.010	0.037	0.019	0.034	0.029
		RMSE	0.086	0.019	0.075	0.019	0.065	0.039
	σ_w=10%	BIAS	0.179	0.038	0.151	0.020	0.131	-0.004
		STD	0.068	0.019	0.065	0.024	0.064	0.035
		RMSE	0.179	0.042	0.166	0.031	0.146	0.035

Table 4. 8 Simulation results for sample size 200

			σ_u= 1%		σ_u= 5%		σ_u= 10%	
			BCC	MVDEA	BCC	MVDEA	BCC	MVDEA
N=500	σ_w= 1%	BIAS	0.015	0.000	0.011	-0.017	0.008	-0.034
		STD	0.008	0.004	0.008	0.014	0.008	0.025
		RMSE	0.017	0.004	0.013	0.022	0.012	0.043
	σ_w= 5%	BIAS	0.088	0.019	0.075	0.001	0.067	-0.020
		STD	0.037	0.010	0.036	0.018	0.035	0.028
		RMSE	0.095	0.022	0.081	0.018	0.076	0.034
	σ_v=10%	BIAS	0.195	0.047	0.168	0.025	0.148	0.002
		STD	0.063	0.018	0.064	0.023	0.064	0.034
		RMSE	0.205	0.050	0.180	0.034	0.161	0.034

Table 4. 9 Simulation results for sample size 500

4.7. Concluding Remarks

A well know limitation of the earlier DEA models, including the standard BCC model, is the sensitivity with respect to outliers and measurement error. However, recently a number of extensions have been developed in the literature that do account for DMU heterogeneity and parameter uncertainty, including outlier detection procedures, robustness measures and stochastic DEA models. These advances can help to increase the robustness of DEA results (efficiency scores, reference units and shadow prices) with respect to outliers and stochastic disturbances.

Mean-Variance Data Envelopment Analyis (MVDEA) is a new framework for incorporating parameter uncertainty and risk aversion in DEA. That framework is consistent with rational choice behavior for general preference structures and general disturbance distributions, and consequently preserves the nonparametric nature of the traditional framework. By contrast, the existing stochastic DEA models invariably require the specification of a particular statistical distribution function for the disturbances.

We have demonstrated that traditional DEA is consistent with stochastic dominance only if the input-output variables are specified accurately or alternatively if the evaluator is risk neutral with respect to parameter uncertainty. If the input-output variables are uncertain and the evaluator is risk averse, the traditional approach may include reference units that are stochastically dominated and exclude units that are stochastically nondominated. By contrast, by using the mean-variance admissible set instead of the traditional set, all stochastically nondominated units are included and all stochastically dominated units are excluded. We have demonstrated how including a variance correction in the objective function can impose the mean-variance set. The empirical application demonstrated that the associated reduction in parameter uncertainty can be substantial.

Future reseach could examine the inclusion of preference information in addition to nonsatiation and risk-aversion (for example, by using third order stochastic dominance rather than second order dominance) and the relaxation of the assumptions used to operationalise SSD (for example, by introducing correlation between the input-output variables).

Apart from selecting stochastically dominated reference units, the traditional model yields efficiency estimates that may suffer from serious bias and variance if the input-output variables are estimated with error. We have demonstrated that variance restrictions can be specified within the mean-variance framework that reduce the uncertainty for the performance of the evaluated unit relative to its reference unit. The simulation results indicate that variance restrictions can improve the efficiency estimates, especially if the sample is large and if parameter uncertainty is important.

Future research could focus on the selection of the appropriate variance restrictions and the effects of such restrictions under a wider range of research environments (including the influence of the production technology and the distribution of observations). Moreover, MVDEA could be compared with alternative techniques, such as those relying on certainty equivalence, chance constained programming, and DEA+.

5. Conclusions and recommendations

5.1. Introduction

The previous chapters have discussed important methodological advances in the DEA methodology. These advances greatly improve the power and the flexibility of DEA to generate valuable information in many different research environments. In addition, a number of novel extensions was developed, using key insights from the field of *Multi-Criteria Decision-Making*, the mathematics of *Generalised Convexity*, and the theory of *Stochastic Dominance*. Section 5.2 summarizes the main conclusions of the previous chapters.

As discussed above, the shortcomings of the original DEA models are fairly well known among economists and econometricians. However, the benefits of DEA as well as the methodological advances that help circumvent the limitations of the earliest models have not been fully appreciated. Therefore, we hope that this study will contribute at the dissemination of DEA, the knowledge of its relative strengths and weaknesses, and the tools currently available for exploiting its full potential.

In addition, many of the methodological advances discussed in this study could be applied in other fields as well. In that respect, the nonparametric analysis of producer analysis forms part of a broader line of research that involves nonparametric tests and measures for the consistency of observed economic behaviour with particular optimising hypotheses. That line of research also includes the nonparametric analysis of consumer behaviour, including Afriat (1967), Diewert (1973), Diewert and Parkan (1978), and Varian (1982). That broader line faces similar problems as DEA: to use the preferences, production and distribution assumptions that match the research environment best. Therefore, the solutions discussed in this study potentially apply to that field as well.

Notwithstanding the recent advances, some problems remain unresolved. Section 5.3 discusses some of the important problems we encountered in this study. The discussion is intended as a marginal note for practical applications of the current DEA methodology. In addition, the remarks are intended as a call for future methodological research.

5.2. Conclusions

Chapter 2 discussed extensions that allow for the analysis of effectiveness, i.e. the extent to which DMUs succeed in achieving their organizational objectives, by including objectives in addition to technical efficiency. Preference information can be included using behavioural assumptions from micro-economic theory, weight restrictions formulated by the decision-maker, and interactive decision procedures. These extensions can greatly improve the goal congruence of DEA models, and improve the relevancy for practical decision-making of DEA results (efficiency measures, reference units and shadow prices).

In addition, Chapter 2 presented *Interactive Data Envelopment Analysis*, a new interactive performance benchmarking procedure that seeks to combine the best of DEA and the interactive decision procedure *Interactive Multiple Goal Programming*. Although the available interactive DEA models circumvent the requirement for a priori articulated general preference information, it is unclear whether

the interactive decision procedures they embed can reliably elicit preference information from decision makers in complex decision problems (i.e. problems with large numbers of decision alternatives and goal variables). Empirical results from experimental psychological research indicate that interactive decision procedures that rely on the information processing operations of selecting a most preferred alternative, or alternatively on assigning criterion weights or weight bounds, are relatively difficult to perform. Such procedures can cause decision-makers to be inconsistent and to use simplified heuristics (e.g. to drop criteria) in complex decision problem, i.e. problems that involve large numbers of decision alternatives and goal variables. By contrast, IMGP relies on the identification of criterion values that must be improved, may deteriorate, or must remain at least equal to the attained satisfactory value. Psychological research indicates that decision-makers can perform that elementary information processing operation with inconsiderable contradictions and complex strategies (e.g. estimating a set of several criteria values).

An additional advantage of using IDEA is that the existing interactive approaches consider efficient frontier units only. By contrast, IMGP allows the decision-maker to select a reference unit from the PPS interior, if, for any reason, he wishes to deliberately include some degree of organisational slack in performance benchmarking.

For these reasons, we believe that the IDEA procedure provides a valuable addition to the tools currently available for including preference information in DEA assessments.

An empirical application for UK university departments illustrated the operation and the characteristics of the IDEA procedure.

Chapter 3 reviewed the advances for imposing and testing production assumptions. A multitude of alternative model formulations exists that allow for a variety of alternative production assumptions. In addition, empirical specification tests can help select the specification that fits the data best. Finally, nonobserved DMUs can be added to the data set to include additional, local production information. These advances greatly extend the flexibility of the DEA methodology to deal with a variety of different production technologies.

Still, most DEA models maintain the assumption that the production possibilities are convex. That assumption simplifies the mathematical structure of the model, because convex problems can be solved with standard mathematical programming techniques. However, it is generally difficult to find a good economic justification for convexity. For example, convexity can be violated by production characteristics such as increasing returns-to-scale, increasing marginal rates of substitution for the inputs and increasing marginal rates of transformation for the outputs. Therefore, it is desirable to develop non-convex models.

Chapter 3 presented such a model, *Transconvex Data Envelopment Analysis*, which uses key insights from the mathematical theory of *Generalised Convexity*. Specifically, TDEA replaces convexity with the more general assumption that the production set is convex transformable, or transconvex, using particular transformations for the inputs and outputs. TDEA embeds some of the important standard models, including the standard BCC model and the less restrictive FDH model.

Sometimes prior economic theory or practical knowledge of the industry under evaluation can provide guidance for the selection of appropriate convexifying transformations. For example, we demonstrated that particular exponential transformations can impose the microeconomic assumption that the input and output sets

are convex. The associated transconvex models have important advantages relative to the model Petersen (1990) proposed for imposing the same aassumption. First, the transconvex models are demonstrably consistent with the assumptions of convexity for the input and/or output sets. Second, the transconvex models allow for the separate introduction and testing of convexity for the input sets and convexity for the output sets. Third, there is no need for the prior specification of the returns-to-scale properties. However, if return-to-scale information is available, it can be incorporated in the transconvex models. Finally, the transconvex models give feasible solutions, whereas the Petersen model may not have a feasible solution for non-increasing returns-to-scale technologies.

In addition, we proposed an empirical selection procedure for selecting the appropriate transformation if the convexifying transformations can not be determined a priori. That procedure allows for a gradual transition between the BCC model and the FDH model, models that represent the limiting cases of the exponential TDEA model. The simulations suggested that the empirical selection procedure suffers from the poor finite sample performance of the embedded specification test. Still, it can substantially reduce bias and error variance relative to the classic models, especially in large samples and if convexity is severely violated.

An empirical application for the bank branch network of a large Dutch financial institution demonstrated that TDEA is not just a theoretical possibility. Theoretical and practical arguments as well as empirical evidence suggested that the bank branch technology is characterised by substantial increasing returns-to-scale. Presumably, this phenomenon can be explained by the limited opportunities small branches have for employing specialist personnel. This finding conflicts with the assumption of convexity imposed by the BCC model, and implies that the BCC model gives inconsistent estimates. In addition, bank branch studies typically involve small or medium sized samples. Therefore, the FDH model can involve substantial finite sample error. However, practical considerations and the data do support convexity of input sets, as is theoretically justified by the law of diminishing marginal rates of technical substitution. Therefore, in contrast to the BCC model, the CIS model does give consistent estimators. In addition, it suffers less from finite sample error than the FDH.

From the theoretical arguments, the simulation results and the practical application, we conclude that the TDEA model is a valuable addition to the research instruments currently available for selecting the production assumptions that best fit the research environment.

Chapter 4 reviewed of the advances presented in the literature for dealing with DMU heterogeneity and parameter uncertainty, including outlier detection procedures, robustness tests, and stochastic DEA models. These developments can help to increase the reliability of DEA results.

In addition, Chapter 4 presented *Mean-Variance Data Envelopment Analysis*, a novel framework for incorporating parameter uncertainty derived from the theory of stochastic dominance. MVDEA considers the selection of reference units in case of parameter uncertainty as a choice problem under uncertainty, and applies the theory of *Stochastic Dominance*, as operationalised by *Mean-Variance Analysis*, to that problem. That theory does not require a specific functional form for the preferences of the decision-maker or a specific distribution function for the statistical distribution of the choice variables. Therefore, Stochastic Dominance it is consistent with the nonparametric nature of the original DEA framework. By contrast, the existing

stochastic DEA models invariably require the specification of a particular statistical distribution function for the disturbances.

We have demonstrated that original DEA framework is consistent with stochastic dominance only if the input-output variables are specified accurately or alternatively if the evaluator is risk neutral with respect to parameter uncertainty. If the input-output variables are uncertain and the evaluator is risk averse, the traditional approach may include reference units that are stochastically dominated and exclude units that are stochastically nondominated. By contrast, by using the mean-variance admissible set instead of the traditional set, all stochastically nondominated units are included and all stochastically dominated units are excluded. We have demonstrated how including a variance correction in the objective function can impose the mean-variance set. An empirical application for large Indonesian commercial banks demonstrated that the associated reduction in parameter uncertainty can be substantial.

Apart from selecting stochastically dominated reference units, the original DEA framework yields efficiency estimates that may suffer from serious bias and variance if the input-output variables are estimated with error. We have demonstrated that variance restrictions can be specified that within the mean-variance framework that reduce uncertainty about the performance of the evaluated unit relative to its reference unit. The results from a Monte-Carlo simulation study indicated that such variance restrictions can substantially improve the efficiency estimates, especially if the sample is large and if parameter uncertainty is important.

From the theoretical arguments, the simulation results and the practical application, we conclude that the MVDEA approach is a potentially valuable addition to the tools currently available to obtain more reliable results than available from deterministic models. Future research could focus on systematically comparing MVDEA with alternative stochastic DEA models, such as those relying on certainty equivalence, chance constrained programming, and DEA+.

5.3. Recommendations

As discussed above, there is much reason for enthusiasm about the methodological advances in DEA. These advances greatly improve the power and the flexibility of DEA to generate valuable information in many different research environments. However, some problems remain unresolved. The problems we encountered in this study include (1) the dynamic and uncertain nature of production as an optimization problem, (2) the goodness of the empirical specification tests currently available, (3) the absence of analytical results on the statistical properties of DEA estimators in case of small samples and noise, and a generally accepted framework for performing Monte-Carlo simulations. In this section, we discuss these problems. The discussion is intended as a marginal note for practical applications of the current DEA methodology. In addition, the remarks are intended as a call for future methodological research.

1. Production a problem of dynamic optimization under uncertainty

DEA views production as a problem of optimization under technological constraints, and measures the extent to which DMUs have succeeded in solving that problem, i.e. efficiency. This approach is consistent with the microeconomic theory of production, and therefore provides a link between empirical production analysis and production theory. However, DEA fails to acknowledge that production in many industries is a problem of optimization *under uncertainty* and a problem of *dynamic* optimization.

The discussion about uncertainty in DEA has focused on the uncertainty that the *analyst* faces in approximating the production possibility set using empirical data. That uncertainty stems from sampling error, i.e. the possibility that the DMUs in the data set do not give a complete representation of the production possibilities (see Section 1.4). Another source of uncertainty for the analyst is parameter uncertainty, i.e. the possibility that the input-output observations are inaccurate estimates for the true input-output variables (see Chapter 4).

A problem not recognized so far is *economic uncertainty*, i.e. uncertainty faced by the DMUs about the outcomes of the production process. For example, in banking, performance generally depends on the uncertain influence of uncontrollable external risk factors such as interest rates, foreign exchange rates and the business cycle. Banks that appear inefficient in DEA may be efficient in reality, if their performance is negatively influenced by such external risk factors.

Economic uncertainty causes problems for the estimation of efficiency that are similar to that caused by parameter uncertainty, because it makes observed performance deviate from the part of performance that is under the control of the DMU. Therefore, the techniques that deal with parameter uncertainty could be helpful in correcting for economic uncertainty as well. However, unlike measurement error, economic risk factors generally can not be modeled as random variables that are independent across DMUs and across risk factors. For example, all banks with a poor asset quality typically are hurt by an economic recession. In addition, economic uncertainty introduces another complication. DEA compares *ex post* production outcomes with *ex post* best practice. *Ex post* outcomes represent the outcomes in a single 'accidental' scenario only, and fail to account for outcomes in alternative scenarios. For example, banks with a poor asset quality can appear efficient in a scenario of economic growth if the loan losses that occur in a scenario of economic recession do not manifest themselves. It is generally desirable to evaluate the complete *ex ante* distribution of outcomes rather than a single scenario only.

In contrast to sampling error, outliers and parameter uncertainty, economic uncertainty has received relatively little attention. DEA studies (including the ones that use stochastic models) invariably represent production as a problem of optimization under certainty. Two interesting approaches to including uncertainty are offered by Spronk and Post (1999b) and Kuosmanen and Post (1999b). Both approaches are based on theories of Stochastic Dominance. As discussed in Chapter 4, Stochastic Dominance provides an interesting approach, because it does not require a specific functional form for the preferences of the decision-maker or a specific statistical distribution function for the statistical distribution of the choice variables, and therefore preserves the nonparametric nature of the original DEA framework.

Post and Spronk (1999b) consider the problems that arise if the inputs and outputs involve uncertainty. They propose to include uncertainty by using decision rules of second-order stochastic dominance, as operationalised by mean-variance analysis[77]. In addition, they propose to represent economic uncertainty by multi-factor risk models. Multi-factor risk models are well-established tools for structuring uncertainty in

---

[77] Chapter 4 of this study used mean-variance analysis to include parameter uncertainty faced by the analyst when evaluating the DMUs. By contrast, Post and Spronk (1999b) use mean-variance analysis to include the economic uncertainty faced by the DMUs under evaluation. A model that accounts for both economic uncertainty and parameter uncertainty is a natural extension. Bawa and Klein (1976, 1977) developed this approach for the problem of portfolio selection.

economic problems. They represent uncertainty as the exposure to unexpected changes in certain stochastic risk factors that are beyond the control of the evaluated units, such as interest rates, foreign exchange rates and the business cycle. These models structure the information required for applying mean-variance analysis, and in addition allow for deriving simplified models that substantially reduce the information requirement and the computational complexity.

Kuosmanen and Post (1999b) consider the problems that arise if the input-output prices faced by the DMUs involve uncertainty. Prices are typically determined by demand and supply forces that are (partly) beyond the control of the DMUs, and in addition that are (partly) unpredictable for the DMUs. In addition, even if the nominal prices would be fixed i.e. by long-term contracts or regulatory acts, the unpredictable fluctuations in uncontrollable environmental factors such as business cycles, interest rates, and inflation guarantee that the real prices in terms of opportunity costs typically involve uncertainties beyond hedging. Thus, evaluating efficiency of DMUs at the realised prices observed *ex post* by the analyst is unfair towards the units who actually did not have all that information at hand when they fixed their production plans. To circumvent this problem, Kuosmanen and Post used the theory of first-order stochastic dominance (Hadar and Russell (1969)) to derive necessary and sufficient conditions for efficiency under uncertainty.

Apart from ignoring the influence of uncertainty on performance, the current DEA model do not acknowledge the dynamic nature of the production process. In many industries the problem of production is dynamic, in the sense that current outputs are the fruits of past inputs in addition to current inputs, and that current inputs yield future outputs in addition to current outputs. For example, in banking, substantial amounts of input can be used to acquire market share or customer goodwill, which generate output in the future in addition to the present. However, DEA typically considered DEA as problem of static optimisation. Therefore, DMUs that appear inefficient may be efficient in reality, if a substantial part of their current inputs are used to generate future outputs.

Only few studies in the DEA literature deal with this problem. Sengupta (1995) presents a dynamic DEA model by introducing the shadow values of quasi-fixed inputs and their optimal paths into an analytic linear programming problem. Färe and Grosskopf (1996) formulate several kinds of intertemporal substitution among inputs, outputs and intermediate inputs using a network theory by which more realistic production processes across periods can be described. Nemoto and Goto (1999) extend DEA to a dynamic DEA framework by incorporating the adjustment costs of investment and intertemporal substitution.

2. Specification tests

As discussed in Chapter 3, the DEA methodology currently includes a wide range of models that impose different production assumptions. Some models impose minimal production assumptions only (e.g. the FDH model, which assumes free disposability only). However, in small samples those models can suffer from substantial finite sample error, which introduces the need to include additional production assumptions (e.g. returns-to-scale properties). Unfortunately, prior theoretical considerations and practical knowledge of the industry under evaluation generally do not suffice to select a particular model specification. Therefore, it is desirable to develop tests that can select the specification based on empirical evidence. In Section 3.6, such an empirical specification test (i.e. an F-test) provided guidance in the selection of the appropriate convexifying

transformation in the TDEA model. The results of the Monte-Carlo simulations suggested that the test could improve estimation substantially.

Unfortunately, the tests currently available are invariably asymptotic by nature, and require a relatively large data set. However, the problem of model selection is relatively unimportant in large samples, as the FDH model can be used in large samples with relatively little finite sample error. Paradoxically, this is reflected in the fact that many tests (including the *F*-test used in Section 3.6) rely on the asymptotic convergence of FDH estimators and more progressive estimators. In addition, many DEA applications (including those presented in this study) involve small samples only, e.g. due to a lack of homogeneous reference units or the proprietary nature of the data. Unfortunately, relatively little is known about the statistical performance of the existing tests in small samples. Moreover, the existing evidence suggests that these tests are sensitive to finite sample error (e.g. Kittelsen (1997)). For example, the results of the Monte-Carlo simulations in Section 3.7 suggested that the empirical selection procedure suffered from the poor finite sample performance of the embedded *F*-test. Therefore, efforts should be directed at gathering systematic evidence on the finite sample performance of existing specification tests, and at designing new ones.

3. Small sample performance and Monte-Carlo simulations

As discussed in Section 1.4, much research has been directed at deriving the asymptotic properties of DEA estimators. However, asymptotic properties relate to performance in large samples only. In addition, the studies currently available ignore the role of noise in addition to sampling error. By contrast, many real-life research environments involve small samples (e.g. for reasons mentioned above) and substantial noise (e.g. because of omitted variables, the use of accounting data rather than economic data, and data-entry error). To the best of our knowledge, there currently are no analytical results on the statistical properties of DEA estimators for research environments that involve small samples and noise. Several studies have studied the statistical properties in case of small samples and noise using Monte-Carlo simulations (including the simulations in Section 3.7 and 4.6). Unfortunately, there currently is no generally accepted framework for performing such studies. For example, it is not clear what production technologies and what statistical distributions for efficiency and noise best describe real-life problems. This is a cause for concern, because the simulation results can have little significance for research practice, if the simulations do not represent a wide range of real-life research environments. Therefore, future research should focus on defining such a framework. That would increase the practical relevance of the simulation results, and in addition make the results of different studies better comparable.

From the above discussion, we conclude that DEA is an attractive methodology. Its strengths include its nonparametric nature and its ability to generate a wealth of information that is specially tailored to individual DMUs (including efficiency measures, reference units and shadow prices). In addition, during the past 20 years a number of methodological advances have greatly increased the possibilities for fully exploiting these strengths in many different research environments. However, some important problems remain unresolved, including the ones discussed above. Therefore, we recommend practitioners to keep these problems in mind when interpreting DEA results. In addition, we recommend combining DEA with alternative techniques for productivity and efficiency analysis, including Econometric Frontier Analysis. Finally, we call for future research to focus on solving the unresolved problems.

References

Adolphson, D. L., G. C. Cornia, and L. C. Walters (1990): 'A Unified Framework for Classifying DEA Models', *Operations Research* 90 (Special edn; H. E. Bradley, ed.): 647-657.

Afriat , S. (1967): 'The construction of a Utility Function from Expenditure Data', *International Economic Review* 8: 67-77.

Afriat , S. (1972): 'Efficiency Estimates of Production Frontiers', *International Economic Review* 13: 568-598.

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

Aigner, D. J., C.A.K. Lovell and P. Schmidt (1977): 'Formulation and estimation of stochastic frontier production models', *Journal of Econometrics* 6 (1): 21-37.

Ali, A. I. (1990): 'Data envelopment Analysis: Computational Issues', *Computers, Environment and Urban Systems* 14 (2): 157-165.

Ali, A. I. (1993): 'Streamlined Computation for Data Envelopment Analysis.' *European Journal of Operational Research* 64 (1).

Ali, A. I. (1994): 'Computational Aspects of Data Envelopment Analysis'. In: Charnes A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (Eds.) (1994*), Data Envelopment Analysis: Theory, methodology and applications,* ISBN 0-7923-9479-8, Kluwer Academic Publishers, Dordrecht, 313-328.

Ali, A. I. and L. M. Seiford (1990): 'Translation Invariance In Data Envelopment Analysis', *Operations Research Letters* 9 (6): 403-405.

Ali, A. I. and L. M. Seiford (1993): 'The Mathematical Programming Approach to Efficiency Measurement.' In: Fried, H., C. A. K. Lovell, and S. Schmidt (eds.) *'The Measurement of Productive Efficiency: Techniques and Applications*.' London: Oxford University Press.

Al-Faraj, T.N., A.S. Alidi and K.A. Bu-Bshait (1993): 'Evaluation of Bank Branches by Means of Data Envelopment Analysis', *International Journal of Operations & Production Management* 13 (8): 45-52.

Allen, R., A. D. Athanassopoulos, R.G. Dyson and E. Thanassoulis (1997): 'Weight Restrictions and Value Jugdements in DEA: Evolution, Development and Future Directions', *Annals of Operations Research* 73: 13-34.

Athanassopoulos, A.D. (1997): 'Service quality and operating efficiency synergies for management control in the provision of financial services: Evidence from Greek bank branches', *European Journal of Operational Research* 98: 301-314.

Atkinson, S. and P.W. Wilson (1995): 'Comparing mean efficiency and productivity scores from small samples: a bootstrap methodology', *Journal of Productivity Analysis* 6: 137-152.

Avriel, M., W. E. Diewert, S. Schaible, and I. Zang (1988): *'Generalised Concavity'*, Plenum Press, New York.

Banker, R.D. (1984): 'Estimating most productive scale size using data envelopment analysis', *European Journal of Operational Research* 17, 35-44.

Banker, R. D. (1993): 'Maximum likelihood, consistency and Data Envelopment analysis: a statistical foundation', *Management Sience*, 1265- 1273.

Banker, R. D. (1996): 'Hypothesis tests using Data Envelopment Analysis', *Journal of Productivity Analysis* 7: 139-159.

Banker, R.D, H.S. Chang and W. W. Cooper (1996): 'Equivalence and implementation of alternative methods for determining returns to scale in DEA', *European Journal of Operational Research* 89: 473-481.

Banker, R.D., Charnes, A. and Cooper W. W. (1984): 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management Science*, 30(9): 1078-1092.

Banker, R.D., Charnes, A. and Cooper W. W. and A. Maindiratta (1988): 'A Comparison of DEA and Translog Estimates of Production Frontier s Using Simulated Observations From a Known Technology', In: Ali Dogramaci, Rolf Färe (Eds.) *'Applications of Modern Production Theory: Efficiency and Productivity'*. Boston: Kluwer Academic Publishers.

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

Banker, R.D., V. M. Gath and W. L. Gorr (1993): 'A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and data envelopment analysis', *European Journal of Operational Research* 67: 332-343.

Banker, R.D. and Maindiratta, A. (1988): 'Nonparametric analysis of technical and allocative efficiencies in production', *Econometrica* 56 (6): 1315-1332.

Banker, R.D. and Morey, R.C. (1986): 'The use of categorical variables in Data Envelopment Analysis', *Management Science* 32 (12): 1613-1627.

Banker, R.D. and R. M. Thrall (1992): 'Estimation of returns to scale using data envelopment analysis', *European Journal of Operational Research* 62: 74-84.

Battese, G. E. and G. S. Corra (1977): 'Estimation of a production frontier model with application to the pastoral zone of eastern Australia', *Australia Journal of Agricultural Economics* 21: 169-179.

Bauer, P. W. (1990): 'Recent developments in the econometric estimation of frontiers', *Journal of Econometrics* 46: 39-56.

Baumol, W. J, J. C. Panzar and R. D. Willig (1982): *'ContesTable Markets and the Theory of Industry Structure'*, Hartcourt Brace Jovanovich, New York.

Bawa, R. W. and V. S. Klein (1976): 'The effect of estimation risk on optimal portfolio choice', *Journal of Financial Economics* 3 (3): 215-231.

Bawa, R. W. and V. S. Klein (1977): 'The effect of limited informationand estimation risk on optimal portfolio diversification', *Journal of Financial Economics* 5 (1): 89-111.

Bazaraa, M. S., H. D. Sherali and C. M. Shetti (1993): *'Non-linear programming'*, ISBN 0-471-55793-5. Second Edition.

Beasley, J. E. (1990): 'Comparing university departments', *Omega*, 171-183.

Belton, V. (1990): 'An integrating Data Envelopment Analysis with Multiple Criteria Decision Analysis', In: *'Multiple Criteria Decision Making'*, Proceedings of the ninth international conference: Theory and applications in business, industry, and government, 71-79.

Belton, V. and Vickers, S. P. (1992): 'VIDEA: Integrated Data Envelopment Analysis and Multiple Criteria Decision Analysis: A visual interactive approach', In: proceedings of the tenth International Conference on Multiple Criteria Decision Making, Vol. II, pp. 419-429.

Belton, V. and Vickers, S. P. (1993): 'Demystifying DEA: A visual interactive approach based on Multi Criteria Analysis', *Journal of the Operational Research Society* 44: 883-896.

Ben-Tal, A. (1977): 'On generalised means and generalised convexity', *J. Optimization Theory Appl.* 21: 1-13.

Berger, A. N. and D. B. Humphrey (1997): 'Efficiency of Financial Institutions: International Survey and Directions for Future Research', *European Journal of Operational Research* 98: 175-212.

Bigelow, J. P. (1993): 'Consistency of mean-variance analysis and expected utility analysis: A complete characterisation', *Economics Letters* 43: 187-192.

Blume, M. and I. Fried (1975): 'The asset structure of individual portfolios and some implications for utility functions', *Journal of Finance* 10: 585-603.

Bogetoft, P. (1996): 'DEA on Relaxed Convexity Assumptions', *Management Science* 42: 457-465.

Bojanic, A. N., S. B. Caudill and J. M. Ford (1998): 'Small-sample properties of ML, COLS, and DEA estimators of frontier models in the presence of heteroscedasticity', *European Journal of operational Research* 108: 140-148.

Bol, G. (1986): 'On Technical Efficiency Measures: A Remark', *Journal of Economic Theory38 (2): 380-385.*

Bol, G. (1988): 'On the definition of Efficiency Measures: A Note', in Eichorn, W., ed. (1988): *'Measurement in Economics: Theory and Applications of Economic Indices,* Physica-Verlag, Heidelberg.

Bowlin, W. F., Charnes, A. and Cooper W. W. and H. D. Sherman (1985): 'Data Envelopment Analysis and Regression Approaches to Efficiency estimation and evaluation.' *Annals of Operation Research* 2 (1): 113-138.

Boyd, G. and R. Färe (1984): 'Measuring the Efficiency of Decision-Making Units: A Comment', *European Journal of Operational Research* 15: 331-332.

Charnes, A., S. Haag, P. Jaska and J. Semple (1992): 'Sensitivity of efficiency classifications in the additive model of Data Envelopment Analysis', *International Journal of Systems Science* 23: 789-798.

Charnes, A. and W. W. Cooper (1985): 'Preface to topics in data envelopment analysis.', *Annals of Operations Research* 2: 59-94.

Charnes, A., W. Cooper, B. Golany, L. Seiford, J. Stutz (1985): 'Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions', *Journal of Econometrics* 30 (1): 91-107.

Charnes A., Cooper, W. W., Huang, Z. M. and Sun, D. B. (1990): 'Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks', *Journal of Econometrics* 46: 73-91.

Charnes A., Cooper, W. W., Lewin, A. Y., R. C. Morey, and J. Rousseau (1985): 'Sensitivity and stability analysis in DEA', *Annals of Operations Research* 2: 139-156.

Charnes A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (Eds.) (1994*): 'Data Envelopment Analysis: Theory, methodology and applications'*, ISBN 0-7923-9479-8, Kluwer Academic Publishers, Dordrecht.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978): 'Measuring the efficiency of decision making units', *European Journal of Operational Research* 2 (6): 429-444.

Charnes, A., Cooper, W. W., Seiford, L. and Stutz, J. (1982): 'A multiplicative model for efficiency analysis', *Socio-Economic Planning Sciences* 16 (5): 223-224.

Charnes, A., Cooper, W. W., Seiford, L. and Stutz, J. (1983): 'Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopments', *Operations Research Letters* 2 (3): 101-103.

Charnes, A. and L. Neralic (1990): 'Sensitivity analysis of the additive model in data envelopment analysis', *European Journal of Operational Research* 48:332-341.

Coelli, T., D. S. P. Rao and G. E. Battese (1998): *'An introduction to efficiency and productivity analysis'*, Kluwer Academic Publishers, ISBN 0-7923-8060-3.

Cohn, R., W. Lewellen, R. Lease and G. Schlarbaum (1975): 'Individual investor riskaversion and investment portfolio composition', *Journal of Finance* 10: 605-620.

Cooper, W.W., Z. M. Huang and S. X. Li (1996): 'Satisficing DEA models under chance constraints'. *The Annals of Operations Research* 66: 279-295.

Cooper, W.W., Z.M. Huang, V. Lelas, S. Xi and O.B. Olesen (1998): 'Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA'. *Journal of Productivity Analysis* 9 (1): 53-79.

Cooper, W. W. and K. Tone (1997): 'Measures of inefficiency in data envelopment analysis and stochastic frontier estimation', *European Journal of Operational Research* 99: 72-88.

Cornwell, C., P. Schmidt and R. C. Sickles (1990): 'Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels', *Journal of Econometrics* 46: 185-200.

Cournot, A. (1838): *Recherches sur les Principes Mathématiques de la Théorie des Richesses.* English edition: *Researches into the Mathematical Principles of the Theory of Wealth,* Edited by N. Bacon, New York, Macmillan.

Debreu, G. (1951): 'The Coefficient of the Resource Utilization', *Econometrica* 19 (3): 273 - 292.

Dekker, D.J. and G.T. Post (1999): 'A quasi-concave DEA model with an application for bank branch performance evaluation', Rotterdam Institute for Business Economics Studies (RIBES) working chapter 9906, ISBN 90-5086-319-1.

Deprins, D., L. Simar and H. Tulkens (1984): 'Measuring labour efficiency in post offices'. In: *'The performance of public enterprises'* (M. Marchand, P. Pestieu and H. Tulkens, eds.), 243-267. North Holland. Amsterdam.

Diewert, W. E. (1973): 'Afriat and Revealed Preference Theory'. *Review of Economic Studies* 40: 419-426.

Diewert, W. E. and Parkan (1978) 'Flexible functional forms and global curvature conditions'. *Econometrica* 25: 43-68.

Diewert, W. E. and Parkan (1983): *'Linear Programming Tests of Regularity Conditions for Production Frontiers'*, in: Eichhorn, W, R. Henn, K. Neumann, and R. W. Shephard, eds. (1983): *'Quantitative Studies on Production and Prices'*, Physica-Verlag, Heidelberg.

Diewert, W. E. and T. J. Wales (1987): 'Flexible functional forms and global curvature conditions', *Econometrica* 25: 43-68.

Drake, L., and B. Howcroft, (1994): 'Relative Efficiency in the Branch Network of a UK Bank: An empirical Study', *Omega : The International Journal of Management Science* 22 (8): 83-90.

Dyson, R. G. and Thanassoulis, E. (1988): 'Reducing weight flexibility in Data Envelopment Analysis', *Journal of the Operational Research Society* 39 (6): 563-576.

Emrouznejad, A.: 'Ali Emrouznejad's DEA HomePage', http://www.csv.warwick.ac.uk/~bsrlu/.

Epstein, M. K. and Henderson, J.C. (1989): 'Data Envelopment Analysis for managerial control and diagnosis', *Decision Sciences* 20: 90-119.

Färe, R., and S. Grosskopf (1996) in *'Intertemporal Production Frontiers: With Dynamic DEA'*, Kluwer, Dordrecht.

Färe, R., and S. Grosskopf (1997): 'Profit efficiency, Farrell decompositions and the Mahler inequality', *Economics Letters* 57: 283-287

Färe, R., S. Grosskopf and C. A. K. Lovell (1985): *'The Measurement of Efficiency of Production'*, Kluwer Academic Publishers.

Färe, R., S. Grosskopf and C. A. K. Lovell (1994): *'Production Frontiers'*, Cambridge University Press.

Färe, R., and W. Hunsaker (1986): 'Notions of Efficiency and Their Reference Sets', *Management Science* 32 (2): 237-243.

Färe, R. and C. A. K. Lovell (1978): 'Measuring the Technical Efficiency of Production', *Journal of Economic Theory* 19 (1): 150-162.

Farrell, M. J. (1957): 'The measurement of efficiency of production', *J. Roy. Statist. Soc*, Ser . A ,120: 251-281.

Ferrier, G. D. and J. G. Hirschberg (1997): 'Bootstrapping confidence intervals for linear programming efficiency scores: With an illustration using Italian bank data', *Journal of Productivity Analysis* 8: 19-33.

Ferrier, G. D. and C. A. K. Lovell (1990): 'Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence', *Journal of Econometrics* 46: 229-245.

Fishburn, P. C. and R. G. Vickson (1978): 'Theoretical Foundations of Stochastic Dominance', In: *'Stochastic Dominance: An Approach to Decision-Making Under Risk'* (G. A. Whitmore and M. C. Findlay, eds.), D.C. Heath, Lexington, MA, ISBN 0-669-013102, 39-114.

Førsund, F. R. and E. S. Jansen (1977): 'On Estimating Average and Best-Practice Homothetic Production Frontiers via Cost Functions', *International Economic Review* 18 (2): 463-476.

Førsund, F. R. and L. Hjalmarsson (1979a): 'Generalised Farrell Measures of Efficiency: An Application to Milk Processing in Swedish Dairy Plants', *Economic Journal* 89 (354): 294-315.

Førsund, F. R. and L. Hjalmarsson (1979a):' Frontier Production functions and technical progress: A study of General Milk Processing in Swedish Dairy Plants' *Econometrica* 47 (4): 883-900.

Gijbels, I., E. Mammen, B.U. Park and L. Simar (1999) : "On Estimation of Monotone and Concave Frontier Functions", *Journal of the American Statistical Association* 194 (445): 220-228.

Golany, B. (1988): 'An interactive MOLP procedure for the extension of DEA to effectiveness analysis', *Journal of the Operational Research Society* 39 (8):725-734.

Golany, B. and Roll, Y. (1994): 'Incorporating standards via DEA'. In: Charnes A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (1994), *'Data Envelopment Analysis: Theory, methodology and applications'*, ISBN 0-7923-9479-8, Kluwer Academic Publishers, Dordrecht, 313-328.

Gong, B.-H. and R. C.Sickles (1989): 'Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data', *Journal of Productivity Analysis* 1: 229-261.

Gong, B.-H. and R. C.Sickles (1992): 'Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data', *Journal of Econometrics* 51: 259-284.

Gong, L. and B. Sun (1995): 'Efficiency measurement of production operations under uncertainty', *International Journal of Production Economics* 39: 55-66.

Green, R. H., J. R. Doyle, and W. D. Cook (1996): 'Efficiency bounds in data envelopment analysis', *European Journal of Operational Research* 89: 482-490.

Grosskopf, S. (1996): 'Statistical inference and nonparametric efficiency: a selective survey', *Journal of Productivity Analysis* 7: 161-176.

Gstach, D. (1995): 'Comparing structural efficiency of unbalanced subsamples: a resampling adaptation of Data Envelopment Analysis', *Empirical Economics* 20: 531-542.

Gstach, D. (1998): 'Another Approach to Data Envelopment Analysis in Noisy Environments: DEA+', *Journal of Productivity Analysis* 9 (2): 161-176.

Hadar, J. and W. R. Russell (1969): 'Rules for Ordering Uncertain Prospects', *American Economic Review* 59: 25-34.

Hall, P., W. Härdle, and L. Simar (1995), "Iterated Bootstrap with Application to Frontier Models", *The Journal of Productivity Analysis* 6: 63-76.

Halme, M., T. Joro, P.Korhonen, S. Salo and J. Wallenius (1999): 'A Value Efficiency Approach to Incorporating Preference Information in Data Envelopmenrt Analysis', *Management Science* 45 (1): 103-115.

Hanoch, G. and H. Levy (1969): 'The efficiency analysis of choices involving risk', *Review of Economic Studies* 36: 335-346.

Hanoch, G. and M. Rothschild (1972): 'Testing the Assumptions of Production Theory: A Nonparametric Approach', *Journal of Political Economy* 80 (2): 256-275.

Hardy, G. H., J. E. Littlewood and G. Golya (1952): '*Inequalities*', 2nd ed., Cambridge University Press, Cambridge, England.

Hay, D. A. and G. S. Liu (1997): 'The Efficiency of Firms: What difference Does Competition Make?', *The Economic Journal* 107: 597-617.

Imperial College Operational Research Library, ftp://graph/ms.ic.ac.uk/pub/.

Jondrow, J., C. A. K. Lovell, I. S. Materov, P. Schmidt (1982): 'On the estimation of technical efficiency in the stochastic frontier production function model', *Journal of Econometrics* 19: 1377-1398.

Kalirajan, K. P. and J. C. Flinn (1983): 'The measurement of farm-specific technical efficiency', *Pakistan Journal of Applied Economics* 2: 167-180.

Kerstens, K. and P. Vanden Eeckaut (1998): 'Distinguishing Technical and Scale Efficiency on Non-Convex and Convex Technologies: Theoretical Analysis and Empirical Illustrations', Louvain-la-Neuve, UCL (CORE Discussion chapter 9855).

Kittelsen, S. A. C. (1997): 'Monte Carlo simulations of DEA efficiency measures and hypothesis tests', research chapter presented at the Fifth European Workshop on Efficiency and Productivity Analysis, October 1997.

Kneip, A., B. U. Park and L. Simar (1998): 'A note on the convergence of nonparametric DEA efficiency measures', *Econometric Theory* 14: 783-793.

Koopmans, T. C. (1951): 'An Analysis of Production as an Efficient Combination of Activities', in: Koopmans, T. C. (ed.): *'Activity Analysis of Production and Allocation'*. Cowles Commision for Research in Economics. Monograph No. 13, John Wiley & Sons Inc., New York.

Koopmans, T. C. (1957): *'Three Essays on the State of Economic Science'*, McGraw Hill study Company, New York.

Korostelev, A.P., L. Simar and A. B. Tsybakov (1995a): 'Efficient estimation of monotone boundaries', *Annals of Statistics* 23(2): 476-489.

Korostelev, A.P., L. Simar and A. B. Tsybakov (1995b): 'On estimation of monotone and convex boundaries', Publications de l'Institut de Statistique de l'Universite de Paris 39, 3-18.

Kuosmanen, T. and G.T. Post (1999a): 'Robust Efficiency Measurement: Dealing with outliers in Data Envelopment Analysis', Rotterdam Institute for Business Economics Studies (RIBES) working chapter 9911, ISBN 90-5086-325-6.

Kuosmanen, T., and G.T. Post (1999b): Nonparametric Efficiency Analysis under Uncertainty: A First-Order Stochastic Dominance Approach, Helsinki School of Economics and Business Administration Working chapters W-230, ISSN 1235-5674, ISBN 951-791-413-X

Kuosmanen, T. and G.T. Post (1999c): 'Measuring Economic Efficiency under Imperfect Price Information: With an Application to European Commercial Banks', Rotterdam Institute for Business Economics Studies (RIBES) working chapter 9934, ISBN 90-5086-346-9.

Land, K. C., A. K. Lovell and S. Thore (1994): 'Chance-Constrained Data Envelopment Analysis', *Managerial and Decision Economics* 14: 541-554.

Larichev, O., Polyakov, O. and Nikiforov, A. (1987): 'Multicriterion linear programming problems (Analytical survey)', *Journal of Economic Psychology* 8: 389-407.

Levy, H. (1992): 'Stochastic dominance and expected utility: survey and analysis', *Management Science* 38 (4): 555-593.

Lewin, A. Y. and R. C. Morey (1981): 'Measuring the Relative Efficiency and Output Potential of Public Sector Organisations: An Application of Data Envelopment Analysis', *International J. Policy Analysis and Info. Systems* 5 (4): 267-285.

Li, S. X. (1998): ' Stochastic models and variable returns to scale in data envelopment analyis', *European Journal of Operational Research 104:* 532-548.

Lötgren, M (1997): *'Essays on efficiency and productivity: Contributions on bootstrap, DEA and stochastic frontier models'*, Stockholm School of Economics, ISBN 91-7258-441-6.

Lovell, C. A. K. and J. T. Pastor (1995): 'Units invariant and translation invariant DEA models', *Operations Research Letters.*

Lovell, C. A. K. and P. Schmidt (1988) 'A comparison of alternative approaches to the measurement of productive efficiency'. In: Dogramaci, A. and R. Färe (eds.) *'Applications of modern production theory: Effciency and production'*, Kluwer Academic Publishers, Boston MA.

Madden, P. (1986): *'Concavity and Optimisation in Microeconomics'*, ISBN 0-631-14192-8.

Meeusen, W. and J. van den Broeck (1977): 'Efficiency estimation from Cobb-Douglas production functions with composite error', *International Economic Review* 18: 435-444.

Nemoto, J. and M. Goto (1999): 'Dynamic data envelopment analysis: modeling intertemporal behavior of a firm in the presence of productive inefficiencies', *Economic Letters* 64: 51-56.

Nerlove, M. (1965): 'Estimation and Identification of Cobb-Douglas Production Functions', Rand McNally, Chicago, Illinois.

Nijkamp, P. and Spronk, J. (1980): 'Interactive Multiple Goal Programming: An evaluation and some results', in: Fandel, G. and T. Gal (eds.)., *'Multiple criteria decision making theory and applications'*, Springer, Berlin etc., 278-293.

Olesen, O. B. (1995): 'Some unresolved problems in data envelopment analysis: A survey', *International Journal of Production Economics* 39: 5-36.

Olesen, O. B. and N. C. Petersen (1995a): 'Chance Constrained Efficiency Evaluation ', *Management Science* 41: 442-457.

Olesen, O. B. and N. C. Petersen (1995b): 'Incorporating quality into data envelopment analysis: a stochastic dominance approach', *Int. J. Production Economics* 39: 117-135.

Olesen, O. B. and N. C. Petersen (1996): 'Indicators of Ill-conditioned Data Sets and Model Misspecification in Data Envelopment Analysis: An Extended Facet Approach', *Management Science* 42 (2), 205-219.

Oral, M., O. Kettani, and R. Yolalan (1992): 'An Empirical Study on Analysing the Productivity of Bank Branches', *I.I.E.Transactions : Industrial Engineering Research and Development* 24 (11): 166-176.

Park, B.U., L. Simar and C. Weiner (1997): 'FDH efficiency scores from a stochastic point of view', Discussion chapter 9715, Institut de Statistique, Universtite Catholique de Louvain, under revision for *Econometric Theory.*

Pastor, J.T., J. L. Ruiz and I. Sirvent (1995): 'A statistical test for nested radial DEA models', Working chapter, Departamento de Estadistica e Investigacion Operativa, Universidad de Alicante, 03071-Alicante, Spain.

Petersen, N. C. (1990): 'Data Envelopment Analysis on a relaxed set of assumptions', *Management Science* 36 (3): 305-314.

Post, G. T. (1997a): 'Performance benchmarking in stochastic environment using Mean-Variance Data Envelopment Analysis', Rotterdam Institute for Business Economics Studies (RIBES) Report R9701/ O, ISBN 90-5086-235-7, under revision for *Operations Research.*

Post, G. T. (1997b): 'Transconcave Data Envelopment Analysis', Rotterdam Institute for Business Economics Studies (RIBES) Report R9721/O, ISBN 90-5086-256-X, under revision for *European Journal of Operational Research.*

Post, G.T. (1998): 'Estimating non-convex production sets using transconcave DEA'. Rotterdam Institute for Business Economics Studies (RIBES) Report, R9847, ISBN 90-5086-303-5.

Post, G. T. and J. Spronk (1999a): 'Performance Benchmarking Using Interactive Data Envelopment Analysis', *European Journal of Operational Research* 115: 472-487.

Post, G. T. and J. Spronk (1999b): 'MFR'. Rotterdam Institute for Business Economics Studies (RIBES) working chapter 9933 ISBN 90-5086-347-7., forthcoming in the book to honor Prof. Po L. Yu, entitled "*New Frontiers of Decision Making for the Information Technology Era*", World Scientific Publishing Co.

Pratt, J. W. (1964): 'Risk aversion in the Small and the Large', *Econometrica* 32: 122-126.

Pulley, L. B. and Y. M. Braunstein (1992): 'A composite cost function with an application to economies of scope in banking', *Review of Economics and Statistics* 74: 221-230.

Räty, T. (1998): *'Efficient Facets, Influential Observations and Revisions of the Best Practice Boundary for DEA'*, Doctoral Dissertation, University of Joensuu, Publications in Social Sciences No: 31.

Rhodes, E. L. (1978) *'Data envelopment Analysis and Approaches for Measuring the Efficiency of Decision-making Units with an Application to Program Follow-Through in U.S. Education.'* Pittsburgh, PA: PH. D. dissertation, School of Urban and Public Affairs, Carnegie-Mellon University.

Rolstaadas, A. (1995): *'Benchmarking: theory and practice'*, Chapman & Hall, London.

Rothschild, M. and J. E. Stiglitz (1970): 'Increasing Risk: A Definition.', *Journal of Economic Theory* 2: 225-243.

Russell, R. R. (1985): 'Measures of Technical Efficiency', *Journal of Economic Theory* 35 (1): 109-126.

Russell, R. R. (1988): 'On the Axiomatic Approach to the Measurement of Technical Efficiency', in Eichorn, W., ed. (1988): '*Measurement in Economics: Theory and Applications of Economic Indices,* Physica-Verlag, Heidelberg.

Russell, R. R. (1990): 'Continuity of Measures of Technical Efficiency', *Journal of Economic Theory* 51 (2): 255-267.

Schaffnit, C., D. Rosen and J. C. Paradi, (1997):.'Best practice analysis of bank branches: An application of DEA in a large Canadian bank', *Operations Research/Management Science* 37: 653-656.

Schmidt, P. (1985-86): 'Frontier Production Functions', *Econometric Reviews* 4 (2):289-328.

Schmidt, P. and R. C. Sickles (1984): 'Production Frontiers and Panel Data', *Journal of Business and Economic Statistics* 2 (4): 367-374.

Seiford, L. M. (1996): 'Data Envelopment Analysis: The Evolution of the State of the Art 1978-1995', *Journal Productivity Analysis* 7: 99-139.

Seiford, L. M. and R. M. Thrall (1990): 'Recent developments in DEA: the mathematical programming approach to frontier analysis', *Journal of Econometrics* 46: 7-38.

Seiford, L. M. and J. Zhu (1998): 'On alternative optimal solutions in the estimation of return to scale in DEA', *European Journal of Operational Research* 108: 149-152.

Sengupta, J. K. (1995) in: *'Dynamics of Data Envelopment Analysis, Theory of Systems Efficiency'*, Kluwer, Dordrecht.

Shephard, R. W. (1953): *'Cost and production functions'*, Princeton University Press, Princeton, NJ.

Shephard, R. W. (1970): *Theory of cost and production functions'*, Princeton University Press, Princeton, NJ.

Sherman H.D. and G. Ladino (1995): 'Managing Bank Productivity Using Data Envelopment Analysis (DEA)', *Interfaces* 25 (14): 60-73.

Simar, L. (1996) : "Aspects of Statistical Analysis in DEA-type Frontier Models", *The Journal of Productivity Analysis* 7: 177-185.

Simar, L. and P. Wilson (1998a): 'Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models', *Management Science* 44 (1): 49-61.

Simar, L. and P. Wilson (1998b): 'A General Methodology for Bootstrapping in Nonparametric Frontier Models', discussion chapter 9811, Institut de Statistique, UCL and CORE.

Simar L. and P. Wilson (1999a): "Statistical Inference in Nonparametric Frontier Models : The State of the Art", Discussion chapter 9904, Institut de Statistique, UCL and CORE, UCL.

Simar, L. and P. Wilson (1999b) : "Estimating and Bootstrapping Malmquist Indices", *European Journal of Operational Research* 115: 459-471.

Simar, L. and P. Wilson (1999c): "Some Problems with the Ferrier/Hirschberg Bootstrap Idea", *Journal of Productivity Analysis* 11: 67-80.

Simar, L. and P. Wilson (1999d): "Of Course We Can Bootstrap DEA Scores ! But Does It Mean Anything ? Logic Trumps Wishful Thinking, *Journal of*

Productivity Analysis 11: 93-97.

Sirvent, I. (1998): 'A Monte Carlo evaluation of some tests of model specification in DEA', Working chapter, Departamento de Estadistica y Matematica Aplicada, Universidad Miguel Hernandez, Elche, 03202-Alicante, Spain.

Spendolini, M. J. (1992): *'The benchmarking study'*, Amion, New York.

Spronk, J. (1981): *'Interactive Multiple Goal Programming: Applications to financial planning'*, Kluwer Nijhoff, Boston.

Stevenson, R. E. (1980): 'Likelihood functions for generalized stochastic frontier estimation', *Journal of Econometrics* 13: 57-66.

Stewart, T. J. (1996): 'Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis', *Journal of the Operational Research Society* 47: 654-665.

Sueyoshi, T. (1990): 'A Special Algorithm for an Additive Model in Data Envelopment Analysis', *Journal of the Operational Research Society* 42 (6): 463-477.

Sueyoshi, T. and Y.-L. Chang (1989): 'Efficient Algorithm for Additive and Multiplicative Models in Data Envelopment Analysis', *Operations Research Letters* 8(4): 205-213.

Thanassoulis E. (1995): 'Assessing police forces in England and Wales using data envelopment analysis.', *European Journal Of Operational Research* 87 (Euro 20th Anniversary Issue.): 641-657.

Thanassoulis, E. and R. Allen (1998): 'Simulating Weight Restrictions in Data Envelopment Analysis by Means of Unobserved DMUs', *Management Science* 44(4): 586-594.

Thompson, R. G., Dharmapala, P. S. and Thrall, R.M. (1995): 'Linked-cone DEA profit ratios and technical efficiency with applications to Illinois coal mines', *International Journal of Production Economics* 39: 99-115.

Thompson, R. G., Langemeier, L. N., Lee, C.-T., Lee, E. and Thrall, R. M. (1990): 'The role of multiplier bounds in efficiency analysis with application to Kansas farming', *Journal of Econometrics* 46: 93-108.

Thompson, R. G., Singleton, Jr. F. D., Smith, B. A. and Thrall, R. M. (1986): 'Comparative site evaluations for locating a high-energy physics lab in Texas'. *Interfaces* 16(6): 35-49.

Tobin, J. (1958): 'Liquidity preferences as behaviour toward risk', *Review of Economic Studies* 25: 65-86.

Tulkens, H. (1993): 'On FDH Efficiency Analysis: Some Methodological Issues and Application to Retail Banking, Courts, and Urban Transit', *The Journal of Productivity Analysis* 4 (1): 183-210.

Varian, H (1982): 'The nonparametric approach to demand analysis', *Econometrica* 50: 945-974.

Varian, H (1984): 'The nonparametric approach to production analysis', *Econometrica* 52: 579-597.

Varian, H (1985): 'Nonparametric Analyis of Optimising Behaviour with Measurement Error', *Journal of Econometrics* 30 (1/2): 445-458.

Von Neumann, J. and O. Morgenstern (1967): *Theory of games and economic behaviour*', Wiley, New York.

Wilson, P. (1995): 'Detecting Influential Observations in Data Envelopment Analysis', *Journal of Productivity Analysis* 6: 27 - 45.

Whitmore, G. A. (1970): 'Third Degree Stochastic Dominance', *American Economic Review* 60: 457-459.

Wong, Y. and Beasley, J. E. (1990): 'Restricting weight flexibility in Data Envelopment Analysis', *Journal of the Operational Research Society* 41: 829-835.

Yeh, Q.-J. (1996): 'The application of Data Envelopment Analyis with financial ratios for bank performance evaluation.', *Journal of the Operational Research Society* 47: 980-988.

Yue, P. (1992): 'Data envelopment analysis and commercial bank performance: a primer with applications for Missouri banks', St. Louis Federal Bank Review.

Zhu, J. (1996): 'Robustness of the efficient DMUs in Data Envelopment Analysis', *European Journal of Operations Research* 90: 451-460.

Zhu, J. and Z. H. Shen (1995): 'A discussion of testing DMUs' returns to scale', *European Journal of Operational Research* 81: 590-596.

Nederlandstalige Samenvatting (Summary in Dutch)

Introductie

Data Envelopment Analysis (verder DEA) is een methodologie voor het bestuderen van productieprocessen en voor het meten van de efficiëntie van productie-organisaties of *Decision-Making Units* (verder DMUs). De basis voor DEA werd gelegd door Charnes, Cooper en Rhodes (1978), Banker, Charnes en Cooper (1984) en Charnes, Cooper, Golany, Seiford en Stutz (1985). De afgelopen 20 jaar is DEA een populaire methodologie geworden, zowel in de wetenschap als in de praktijk van het bedrijfsleven en van overheidsinstellingen. Van deze populariteit getuigt onder andere een snel groeiend aantal publicaties in vooraanstaande wetenschappelijke toptijdschriften, waaronder *Econometrica*, *Management Science*, en *Operations Research*.

DEA maakt deel uit van een algemenere onderzoeksmethodologie die bekend staat als *Efficient Frontier Analysis* (verder EFA). Dat is een methodologie voor empirische productie-analyse die is gebaseerd op de micro-economische productietheorie. Daarin wordt productie beschouwd als een optimalisatieprobleem onder technologische randvoorwaarden. Die randvoorwaarden worden weergegeven door een *efficiënte productiegrenslijn*. Inefficiëntie is gedefinieerd als de afwijking van gerealiseerde prestaties van de 'optimale' prestaties gegeven de grenslijn. Helaas kan deze productietheorie niet direct worden toegepast in de praktijk, omdat de informatie over de productiemogelijkheden doorgaans onvolledig is. EFA construeert een *empirische grenslijn* verkregen op basis van een verzameling productiegegevens van vergelijkbare DMUs, en meet de efficiëntie van de DMUs ten opzichte van die empirische grenslijn. De basis voor deze methodologie is in de jaren '50 gelegd door Debreu (1951), Koopmans (1951, 1957) en Farrell (1957).

DEA vertegenwoordigt de *non-parametrische* benadering van EFA. Een andere benadering staat bekend als *Econometric Frontier Estimation* (verder EFE). EFE vertegenwoordigt de *parametrische* benadering van EFA. De basis voor deze methodologie is gelegd door Aigner, Lovell and Schmidt (1977), Meeusen and Van den Broeck (1977) and Battese and Corra (1977). De econometrische benadering schat grenslijnen door middel van regressieanalyse. Eerst wordt een regressievergelijking gespecificeerd die de functionele vorm van de grenslijn (bijvoorbeeld een lineaire of een log-lineaire vorm) vastlegt, maar niet de parameters. De parameters worden vervolgens met optimalisatiemethoden (bijvoorbeeld de methode van de kleinste kwadraten of de methode van de hoogste aannemelijkheid) zo gekozen dat de regressielijn de gegevensverzameling het beste beschrijft. Tenslotte worden de inefficiënties geschat met behulp van de conditionele kansverdeling van de residuen, gegeven een statistische verdelingsfunctie voor de inefficiënties. Een beperking van deze methodologie is dat de conclusies afhankelijk zijn van de specificatie van de functionele vorm van de grenslijn en een bepaalde statistische verdelingsfunctie voor de inefficiënties. Dit is met name een probleem als er geen goede theoretische argumenten bestaan voor een bepaalde specificatie en als empirische specificatietoetsen geen uitsluitsel kunnen geven. Een andere beperking is dat EFE zich vooral richt op het beschrijven van de algemene structuur van de productietechnologie. In management toepassingen is het vaak wenselijk om specifieke informatie te vinden over individuele DMUs. Bovendien is het vaak

wenselijk om (naast het meten van inefficiënties) verklaringen voor inefficiënties te vinden en remedies te bepalen.

DEA biedt een mogelijke oplossing voor deze problemen. Het vereist namelijk geen expliciete functionele vorm voor de grenslijn of een expliciete statistische verdelingsfunctie voor inefficiënties. In plaats van een expliciete specificatie berust DEA op bepaalde algemene vooronderstellingen over de productiestructuur en de statistische verdeling van de observaties.

Bovendien levert DEA een veelheid aan informatie die specifiek gericht is op individuele DMUs. Die informatie omvat onder andere efficiëntiemaatstaven, referentie-eenheden en schaduwprijzen. De *efficiëntiemaatstaven* meten de efficiëntie in het verleden en de mogelijke verbeteringen voor de toekomst. De *referentie-eenheden* zijn een unieke eigenschap van DEA. De referentie-eenheden zijn potentiële *benchmark partners*. Door de bedrijfsvoering van een inefficiënte eenheid te vergelijken met die van zijn *benchmark partners*, kunnen verklaringen worden gevonden en remedies bepaald. De *schaduwprijzen* (of *virtuele gewichten*) weerspiegelen de lokale schaaleffecten, substitutiemogelijkheden van de input variabelen en de transformatiemogelijkheden van de output variabelen.

Ondanks de non-parametrische aard en de rijkdom aan DMU-specifieke informatie, had de DEA methodologie aanvankelijk duidelijke beperkingen. De oorspronkelijke modellen maken namelijk vooronderstellingen die in bepaalde onderzoekssituaties beperkend kunnen zijn. Bovendien was aanvankelijk niet duidelijk hoe die vooronderstellingen getoetst kunnen worden en hoe ze eventueel gewijzigd kunnen worden.

In dit boek onderscheiden we drie soorten vooronderstellingen. Ten eerste maken DEA modellen vooronderstellingen over de organisatorische doelstellingen van de DMUs (verder: preferentievooronderstellingen). Ten tweede worden vooronderstellingen gemaakt over de aard van de productietechnologie (verder: productievooronderstellingen). Ten derde maakt DEA vooronderstellingen over de statistische verdeling van de waarnemingen in de gegevensverzameling (verder: distributievooronderstellingen).

Het is in het algemeen onmogelijk om een model te maken dat geen vooronderstellingen maakt. Preferentievooronderstellingen zijn nodig voor de relevantie van de resultaten (efficiëntiemaatstaven, referentie-eenheden en schaduwprijzen) voor de beleidsvorming. Prestatiemeting is namelijk onmogelijk zonder de beleidsdoestellingen te kennen. Productievooronderstellingen zijn nodig om de empirische grenslijn te construeren op basis van de gegevensverzameling. Sommige modellen maken alleen algemene vooronderstellingen. Helaas werken modellen met algemene vooronderstellingen doorgaans alleen goed als er een relatief grote steekproef is. In kleine steekproeven lijden deze modellen aan *finite sample error*. Distributievooronderstellingen zijn nodig omdat de empirische grenslijn wordt geconstrueerd op basis van de gegevensverzameling en omdat DMUs worden beoordeeld op basis van hun waargenomen prestaties.

Modellen met minimale vooronderstellingen kunnen dus leiden tot irrelevante of onbetrouwbare resultaten. Daar staat tegenover dat modellen met foutieve vooronderstellingen kunnen lijden onder specificatiefouten. Zo kan het oorspronkelijke model van Charnes, Cooper an Rhodes (1978), dat constante schaaleffecten verondersteld, in geval van schaalvoordelen de efficiëntie van kleine (grote) DMUs systematisch onderschatten (overschatten). De resulterende afweging zal verschillen van de ene onderzoekssituatie tot de andere. Het is daarom moeilijk om één 'optimaal' model te ontwikkelen. Het is echter wel wenselijk om te beschikken

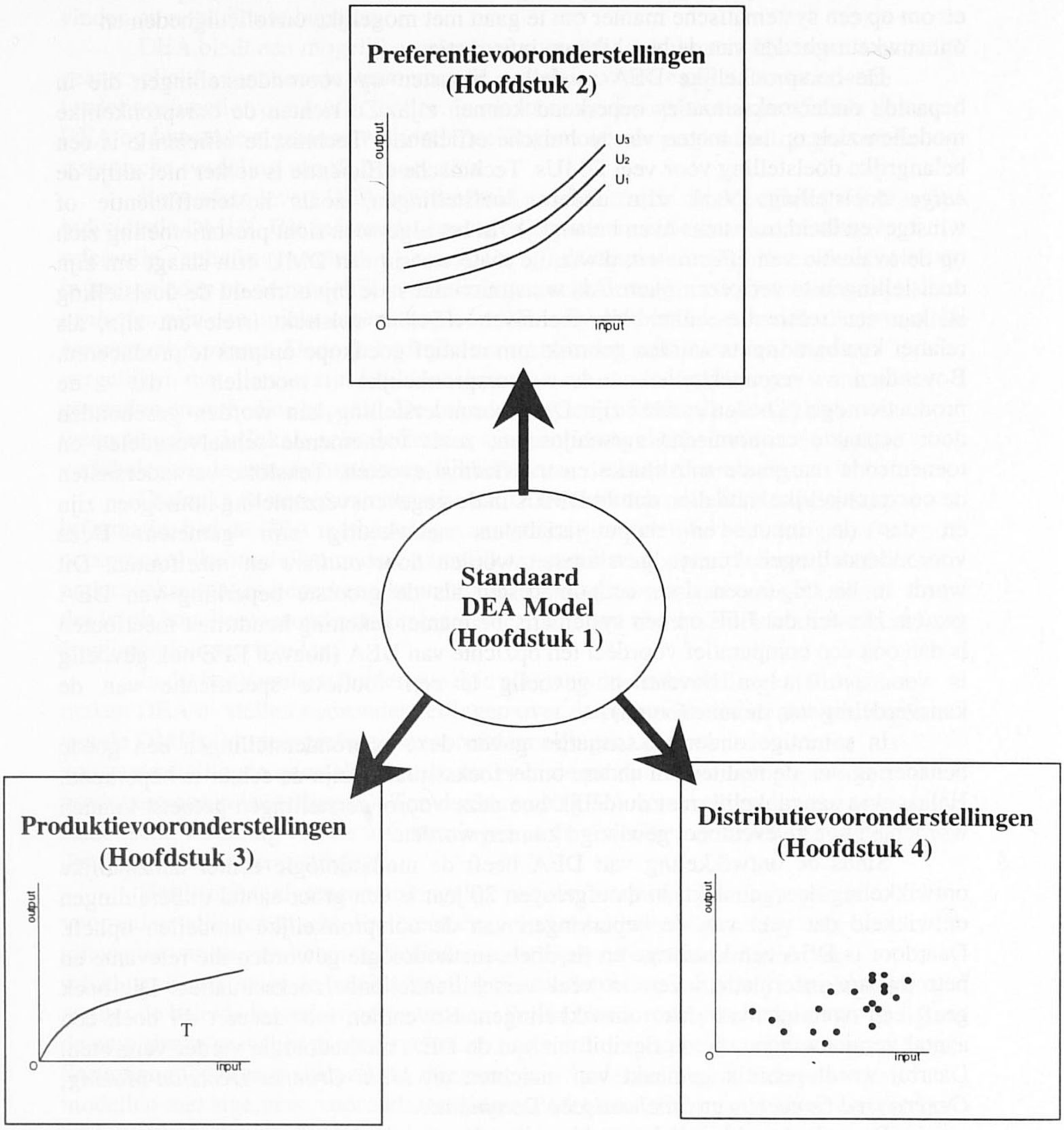
over een methodologie die flexibel genoeg is om beschikbare informatie op te nemen en om op een systematische manier om te gaan met mogelijke onvolledigheden en onnauwkeurigheden van de beschikbare informatie.

De oorspronkelijke DEA modellen berusten op vooronderstellingen die in bepaalde onderzoekssituaties beperkend kunnen zijn. Zo richten de oorspronkelijke modellen zich op het meten van technische efficiëntie. Technische efficiëntie is een belangrijke doelstelling voor veel DMUs. Technische efficiëntie is echter niet altijd de *enige* doelstelling. Vaak zijn andere doelstellingen, zoals kostenefficiëntie of winstgevendheid, minstens even belangrijk. In het algemeen richt prestatiemeting zich op de evaluatie van e*ffectiviteit*, d.w.z. de mate waarin een DMU erin slaagt om zijn doelstellingen te verwezenlijken. Als winstmaximalisatie bijvoorbeeld de doelstelling is, kan een referentie-eenheid die technisch efficiënt volstrekt irrelevant zijn, als relatief kostbare inputs worden gebruikt om relatief goedkope outputs te produceren. Bovendien veronderstellen de oorspronkelijke modellen dat de productiemogelijkheden convex zijn. Die vooronderstelling kan worden geschonden door bepaalde economische verschijnselen, zoals toenemende schaalvoordelen en toenemende marginale substitutie- en transformatievoeten. Tenslotte veronderstellen de oorspronkelijke modellen dat de DMUs in de gegevensverzameling homogeen zijn en dat de input- en output-variabelen nauwkeurig zijn gemeten. Deze vooronderstellingen kunnen geschonden worden door *outliers* en meetfouten. Dit wordt in het algemeen door econometristen als de grootste beperking van DEA gezien. Het feit dat EFE op een systematische manier rekening houdt met meetfouten is dan ook een comparatief voordeel ten opzichte van DEA (hoewel EFE ook gevoelig is voor *outliers* en bovendien gevoelig is een foutieve specificatie van de kansverdeling van de meetfouten).

In sommige onderzoekssituaties geven deze vooronderstellingen een goede benadering van de realiteit. In andere onderzoekssituaties zijn ze echter te beperkend. Helaas was aanvankelijk niet duidelijk hoe deze vooronderstellingen getoetst kunnen worden en hoe ze eventueel gewijzigd kunnen worden.

Sinds de ontwikkeling van DEA heeft de methodologie echter aanzienlijke ontwikkeling doorgemaakt. In de afgelopen 20 jaar is een groot aantal uitbereidingen ontwikkeld dat veel van de beperkingen van de oorspronkelijke modellen opheft. Daardoor is DEA een krachtige en flexibele methodologie geworden die relevante en betrouwbare informatie levert in veel verschillende onderzoekssituaties. Dit boek geeft een overzicht van deze ontwikkelingen. Bovendien introduceert dit boek een aantal vernieuwingen die de flexibiliteit van de DEA methodologie verder vergroten. Daarbij wordt gebruik gemaakt van inzichten uit *Multi-Criteria Decision-Making*, *Generalised Convexity* en *Stochastische Dominantie*.

De rode draad bij de bespreking van deze methodologische vernieuwingen is bovengenoemde driedeling: preferentie-, productie-, en distributie-vooronderstellingen. Figuur 1 toont de structuur van dit boek. Hoofdstuk 1 bespreekt de structuur en de eigenschappen van het standaard DEA model. Hoofdstuk 2 behandelt de uitbreidingen met betrekking tot preferentievooronderstellingen (in figuur 1 afgebeeld als de indifferentiecurven van een hypothetische nutsfunctie (U_1, U_2 and U_3)). Hoofdstuk 3 behandelt productievooronderstellingen (weergegeven door de hypothetische productiemogelijkhedenverzameling T). Tenslotte bespreekt hoofdstuk 4 distributievooronderstellingen (afgebeeld door een hypothetische gegevensverzameling). De volgende paragrafen geven een korte samenvatting van de inhoud van deze afzonderlijke hoofdstukken.



Figuur 1 Hoofdstructuur van dit boek

Hoofdstuk 1: Het standaard DEA model

Hoofstuk 1 is de basis voor de volgende hoofdstukken. Het behandelt de elementaire concepten en eigenschappen van de DEA methodologie en introduceert de notatie die in dit boek wordt gehanteerd. Voor het gemak beperkt hoofdstuk 1 zich tot de zogenaamde output-georiënteerde versie van het standaard model van Banker, Charnes en Cooper (1984).

Het BCC model is het standaard model van de DEA methodologie. Een voordeel van dat model is dat het, in tegenstelling tot het oorspronkelijke Charnes, Cooper and Rhodes (1978) model, de vooronderstelling van constante schaaleffecten loslaat. De meeste methodologische artikelen richten zich op het uitbreiden van het BCC model en de meeste toepassingen gebruiken dat model. Bovendien is het model in de meeste software-paketten opgenomen.

Het BCC model maakt echter bepaalde vooronderstellingen omtrent de beleidsdoelstellingen, de productiemogelijkheden en de kansverdeling van de waarnemingen. Die vooronderstellingen zijn geschikt voor bepaalde onderzoekssituaties, maar kunnen te beperkend zijn voor andere onderzoekssituaties, met name als: (1) er andere doestellingen zijn dan technische efficiëntie, (2) als de productiemogelijkheden niet convex zijn, en (3) als de gegevensverzameling is vervuild door *outliers* of meetfouten. De volgende hoofdstukken bespreken de methodologische uitbereidingen van DEA in termen van het output-georiënteerde BCC model. We zijn er echter van overtuigd dat de uitbreidingen evenzeer van toepassing zijn op alternatieve modellen.

Voor niet-experts is het lezen van hoofdstuk 1 een voorwaarde om die uitbreidingen te begrijpen. De volgende hoofdstukken laten zich evenwel onafhankelijk van elkaar lezen. Daardoor kan elke lezer de meest efficiënte route naar een specifiek onderwerp kiezen.

Hoofdstuk 2: Preferentievooronderstellingen

In de DEA gemeenschap is de noodzaak onderkend om naast technische efficiëntie aanvullende doelstellingen op te nemen. Zonder doelcongruentie zijn de resultaten (efficiëntiemaatstaven, referentie-eenheden en schaduwprijzen) immers irrelevant voor de beleidsvoering. Een aantal benaderingen is voorgesteld. Daarbij gaat het er niet zozeer om welke preferentievooronderstellingen het beste zijn, maar vooral hoe preferentie-informatie kan worden verkregen en hoe de beschikbare informatie kan worden opgenomen in het model. We onderscheiden drie soorten benaderingen.

Allereerst zijn modellen ontwikkeld op basis van gedragsvooronderstellingen uit de micro-economie, zoals kostenminimalisatie, opbrengstmaximalisatie en winstmaximalisatie. Die modellen maken gebruik van relatieve prijzen om bepaalde input en output variabelen te aggregeren. Een beperking is dat prijzen vaak onvolledig of onzeker zijn.

Daarnaast kunnen restricties worden opgenomen die door de besluitvormer zelf zijn geformuleerd. Met name restricties van de variabelen van het duale model (de schaduwprijzen of virtuele gewichten) zijn veel toegepast. Een beperking is dat vooraf veel algemene preferentieinformatie moet worden verzameld. Daarbij kan de besluitvormer worden geconfronteerd met hypothetische keuzes en afwegingen die voor hem weinig herkenbaar zijn.

Tenslotte kan DEA worden gecombineerd met interactieve beslissingsmethoden. Interactieve methoden zijn aantrekkelijk omdat ze de noodzaak tot het op voorhand specificeren van algemene preferentieinformatie verminderen. Ze vertrouwen namelijk

op de interactieve formulering van lokale preferentie-informatie die los komt als de besluitvormer wordt geconfronteerd met bepaalde goed gedefinieerde en haalbare alternatieven.

Het effect van deze vernieuwingen is een hogere doelcongruentie van de modellen en daardoor een hogere beleidsrelevantie van de resultaten. Deze uitbreidingen geven DEA een comparatief voordeel ten opzichte van de econometrische benadering, die minder flexibel is in het opnemen van preferentieinformatie[78].

Interactive Data Envelopment Analysis (Post and Spronk (1999a)), verder IDEA, is een nieuwe interactieve methode. Het is onduidelijk of de bestaande interactieve methoden wel betrouwbaar zijn, vooral als er veel efficiënte referentie-eenheden zijn en als het aantal input en output variabelen groot is. De bestaande methoden berusten namelijk op het selecteren van een meest geprefereerd alternatief, of het toewijzen van gewichten of gewichtsrestricties. Experimenteel psychologisch onderzoek (b.v. Larichev *et al.* (1987)) wijst er op dat deze informatieverwerkende activiteiten relatief moeilijk zijn. Het blijkt namelijk dat veel mensen inconsistent zijn en vereenvoudigende beslissingsregels hanteren bij deze activiteiten, vooral als het aantal alternatieven en doelvariabelen toeneemt.

IDEA maakt gebruik van *Interactive Multiple Goal Programming* (Spronk (1981)). Die interactieve methode is gebaseerd op het aanwijzen van attribuutwaarden (i.g.v. DEA: input en output variabelen) die verbeterd moeten worden, die mogen verslechteren, of die tenminste het behaalde niveau moeten behouden. Psychologisch onderzoek wijst erop dat mensen deze elementaire informatieverwerkende handeling relatief gemakkelijk kunnen uitvoeren zonder tegenstrijdigheden en met complexe strategieën.

Een ander argument voor IDEA is dat de bestaande methoden alleen *efficiënte* referentie-eenheden beschouwen. Zoals hierboven besproken, is technische efficiëntie vaak een doelstelling en vereenvoudigt deze restrictie het keuzeprobleem. Soms is het echter wenselijk om juist *inefficiënte* referentie-eenheden te kiezen. Daarvoor kunnen verschillende psychologische en organisatorische redenen bestaan. Het kan bijvoorbeeld demoraliserend werken als een *benchmark* wordt gehanteerd die niet kan worden overtroffen (hetgeen bij efficiënte referentie-eenheden per definitie het geval is).

Ter illustratie van IDEA is een toepassing opgenomen waarin de onderzoeks- en onderwijsprestaties van de faculteiten van Britse universiteiten wordt gemeten. Daarbij wordt gebruik gemaakt van een gegevensverzameling die eerder door Beasley (1990) is gebruikt voor een analyse waarin preferentieinformatie werd opgenomen met behulp van gewichtsrestricties.

Hoofdstuk 3: Productievooronderstellingen

De DEA methodologie omvat een veelheid aan alternatieve modellen die verschillende productievooronderstellingen maken. Bovendien zijn empirische specificatietoetsen ontwikkeld die kunnen helpen de specificatie te kiezen die de waarnemingen in de gegevensverzameling het beste beschrijft. Tenslotte bestaat er de mogelijkheid om virtuele waarnemingen te introduceren om aanvullende lokale productie-informatie op te

---

[78] In de econometrische benadering kunnen micro-economische gedragsvooronderstellingen worden opgenomen door een kosten-, opbrengsten- of winstgrenslijn te schatten in plaats van een productiegrenslijn. Er bestaan echter geen mogelijkheden om hiervan af te wijken door (zoals in DEA) preferentie-informatie van een specifieke besluitvormer op te nemen.

nemen. Deze instrumenten vergoten de flexibiliteit van DEA en kunnen de betrouwbaarheid van de resultaten aanzienlijk verhogen.

De meeste modellen (waaronder het standaard BCC model) veronderstellen echter dat de productiemogelijkheden convex zijn. Die vooronderstelling vereenvoudigt de wiskundige structuur van het model, omdat convexe problemen relatief eenvoudig oplosbaar zijn. Het is echter moeilijk een goede economische rechtvaardiging te vinden voor convexiteit. Convexiteit is bijvoorbeeld strijdig met economische verschijnselen als toenemende schaalvoordelen, toenemende marginale transformatievoeten voor de output variabelen en toenemende marginale substitutievoeten voor de input variabelen. Daarom is het wenselijk om non-convexe modellen te ontwikkelen. Het *Free Disposable Hull* (verder FDH) model (Deprins, Simar en Tulkens 1984) is zo'n non-convex model. Het maakt alleen de algemene micro-economische vooronderstelling dat de marginale productiviteit van de input variabelen positief is. De FDH schatters zijn statistisch consistent (asymptotisch zuiver en met een verdwijnende variantie) voor algemene productiestructuren. Helaas werkt dat model alleen als er een relatief grote steekproef is. In kleine steekproeven lijdt het model aan *finite sample error*. Bovendien heeft het een relatief lage convergentiesnelheid. Daarom is het wenselijk non-convexe modellen te ontwikkelen die beter presteren in kleine steekproeven.

Het *Transconvex Data Envelopment Analysis* model (Post (1997b, 1999), Dekker and Post (1999)) is zo'n model. Het maakt gebruik van inzichten uit de wiskundige theorie van *Generalised Convexity*. TDEA vervangt de convexiteitsvooronderstelling door de vooronderstelling dat de productiemogelijkheden convex transformeerbaar (of transconvex) zijn met behulp van bepaalde transformaties van de input- en output-variabelen. Het TDEA model omvat enkele belangrijke bestaande DEA modellen, waaronder het BCC model en het FDH model.

Hoewel het TDEA model non-convexe productiemogelijkheden toelaat, blijft het model zelf convex. Daardoor kan het met standaard optimalisatiemethoden worden opgelost.

De geschikte transformaties kunnen soms worden gekozen met behulp van economische theorie. Een mogelijkheid is bijvoorbeeld het gebruik van de algemene micro-economische vooronderstelling dat de *isoquanten* van de productieverzameling convex zijn, hetgeen een minder restrictieve vooronderstelling is dan convexiteit voor de gehele productieverzameling. Deze vooronderstelling werd eerder door Petersen (1990) voorgesteld. Het model dat Petersen daarvoor voorstelde lijdt echter aantoonbaar aan een aantal belangrijke beperkingen. Bepaalde exponentiële transformaties kunnen die vooronderstelling echter opleggen in het TDEA model.

Als economische theorie tekort schiet, kan gebruik gemaakt worden van empirische specificatie toetsen. We ontwikkelen zo'n toets op basis van de F-toets van Banker (1993). Die specificatietoets laat een geleidelijke overgang van het FDH naar het BCC model toe.

De resultaten van Monte-Carlo simulaties wijzen erop dat de efficiëntieschatters aanzienlijk kunnen verbeteren ten opzichte van het BCC model, zeker als de gegevensverzameling groot is en er sprake is van aanzienlijke schendingen van de convexiteitsvooronderstelling. Bovenden is de verbetering ten opzichte van het FDH aanzienlijk, zeker als de gegevensverzameling klein is en een relatief zwakke transformatie volstaat.

Een toepassing voor het kantorennetwerk van een grote Nederlandse financiële instelling laat zien dat schendingen van convexiteit niet alleen een theoretische mogelijkheid zijn. Er zijn namelijk aanwijzingen voor toenemende schaalvoordelen voor de kantoren van die instelling. Daardoor wordt de convexiteitsvooronderstelling

geschonden en is het BCC fout gespecificeerd. Bovendien is de steekproef relatief klein, waardoor het FDH een lage onderscheidende kracht kan hebben. De vooronderstelling dat de isoquanten convex zijn lijkt echter wel gerechtvaardigd, als gevolg van de afnemende marginale substitutievoet voor verschillende soorten kantoorpersoneel (*back officers* en *front officers*). Daarom is een TDEA model dat convexe isoquanten veronderstelt een mogelijk alternatief voor BCC en FDH. Vergelijking met de uitkomsten van een econometrisch model wijzen er op dat TDEA de schatters inderdaad verbetert.

Hoofdstuk 4: Distributievooronderstellingen

In de DEA literatuur is een aantal uitbreidingen ontwikkeld die rekening houden met *outliers* en meetfouten. Zo zijn er methodes ontwikkeld voor het opsporen van *outliers* in DEA en voor het meten van de robuustheid van de resultaten voor verstoringen van de gegevensverzameling. Bovendien zijn er stochastische DEA modellen ontwikkeld die expliciet rekening houden met de mogelijkheid dat de gegevensverzameling is verstoord door meetfouten. Veel van die uitbreidingen zijn gebaseerd op *Stochastic Programming*. Deze uitbreidingen verhogen de betrouwbaarheid van de resultaten. Ze verhelpen het belangrijkste nadeel van DEA ten opzichte van econometrische modellen.

Mean-Variance Data Envelopment Analysis (Post 1997a), verder MVDEA, is een nieuwe stochastische methode. MVDEA beschouwt de keuze van een referentie-eenheid in het geval van *parameter uncertainty* als een keuzeprobleem onder onzekerheid en past de theorie van *Stochastiche Dominantie* toe op dat probleem. De theorie van Stochastische Dominantie is een theorie over besluitvorming onder onzekerheid die de afgelopen decennia sterk is ontwikkeld en veel is toegepast op diverse economische problemen. Die theorie vereist geen expliciete functionele vorm voor de preferentiestructuur van de besluitvormer en geen expliciete distributiefunctie voor de statistische verdeling van de keuzevariabelen. Daardoor sluit het goed aan bij het nonparametrische karakter van DEA. De bestaande stochastische DEA modellen vereisen daarentegen zonder uitzondering de expliciete specificatie van een statistische verdeling van de variabelen.

De oorspronkelijke DEA modellen zijn alleen consistent zijn met stochastische dominantie als de input en output variabelen juist zijn gespecificeerd of als de besluitvormer risiconeutraal is met betrekking tot *parameter uncertainty*. Als de variabelen echter onzeker zijn en de beslisser risico-avers is, dan kunnen de oorspronkelijke modellen stochastisch gedomineerde referentie-eenheden selecteren en stochastisch ongedomineerde eenheden uitsluiten. We laten zien dat bepaalde variantiecorrecties dit probleem kunnen voorkomen. Een empirische toepassing voor grote Indonesische banken toont aan dat de reductie in *parameter uncertainty* aanzienlijk kan zijn. Bovendien laten we zien dat bepaalde variantierestricties de DEA efficiëntieschatters aanzienlijk kunnen verbeteren. De resultaten van Monte-Carlo simulaties tonen aan dat de verbeteringen aanzienlijk kunnen zijn, met name als een grote steekproef beschikbaar is en als er relatief veel onzekerheid is.

Ondanks de methodologische ontwikkelingen blijven bepaalde problemen onopgelost. Paragraaf 5.3 besluit deze studie met een korte bespreking van de onopgeloste problemen die we bij het uitvoeren van deze studie zijn tegengekomen. De bespreking is een belangrijke kanttekening bij praktische toepassingen van de bestaande DEA

methodologie. Bovendien vormt het een mogelijke agenda voor toekomstig methodologisch onderzoek.

The Tinbergen Institute is the Netherlands Research Institute and Graduate School of Economics, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University in Rotterdam, the University of Amsterdam and the Free University in Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. If available, trade editions of the books which are published in the Tinbergen Institute Research Series can be ordered through Thela Thesis, Prinseneiland 305, 1013 LP Amsterdam, the Netherlands, phone: +3120 6255429; fax: +3120 6203395. The following books recently appeared in this series:

160. P.A. GAUTIER, *The flow approach to the labor market.*
161. A.A.R.J.M. VAN RIXTEL, *Informality, bureaucratic control and monetary policy: The case of Japan.*
162. Y. GAO, *Management support of information system project planning.*
163. J.G. PIEK, *Development and acquisition of technology in rural collective industries. The case of China's hinterland.*
164. F. LANGERAK, *The effect of market orientation on the performance of industrial firms: A study in the Dutch Metalelektro.*
165. A.K. SINHA, *Estimating failure probability when failure is rare: multi-dimensional case.*
166. J.P.M. SIJM, *Food security and policy interventions in sub-Saharan Africa. Lessons from the past two decades.*
167. W.A. LETTERIE, *Partisan politics, persuasion and information.*
168. R. VAN DIJK, *Corporate finance policy and equity investment: panel data analyses.*
169. M.A. CARREE, *Market dynamics, evolution and smallness.*
170. J.E. LIGTHART, *Environment, imperfect markets, and public finance.*
171. R. PAAP, *Markov trends in macroeconomic time series.*
172. J.H. ABBRING, *Essays in labour economics.*
173. J.P. COHEN, *Utility: a real thing. A study of utility's ontological status.*
174. G.A.R. VAN HOESEL, *Beyond export-led growth: the emergence of new multinational enterprises from Korea and Taiwan.*
175. H.H.J. ZEBREGS, *International capital movements and technology in economic development.*
176. M.F. CORNET, *Game-theoretic models of bargaining and externalities.*
177. M.N. BOUMAN, *Environmental costs and capital flight.*
178. L. PENG, *Second order condition and extreme value theory.*
179. P.W.C. KONING, *Structural empirical analysis of labor market frictions, job search and wage formation.*
180. C.W.M. NAASTEPAD, *The public sector budget and macroeconomic performance: A real-financial CGE analysis with portfolio choice with reference to India.*
181. S.T.M. STRAETMANS, *Extreme financial returns and their comovements.*
182. J.P. KOOIMAN, *Topics in the Economics of Environmental Regulation.*
183. P.P.A.A.H. KANDELAARS, *Material-product chains: Economic models and applications.*

184. N. VAN GIERSBERGEN, *Bootstrapping dynamic econometric models.*
185. K.M. SCHILSTRA, *Industrial relations and human resource management: A network approach.*
186. H. SUPRATIKNO, *Competitive strategy, production organization and inter-firm collaboration. A case study of subcontracting arrangements in three manufacturing firms in Indonesia.*
187. J. GOUWENS, *Interne consistentie in personeelsmanagement: Naar een typering van HRM in grote bedrijven in Nederland.*
188. F. POT, *Continuity and change of human resource management: A comparative analysis of the impact of global change and cultural continuity on the management of labour between the Netherlands and the United States.*
189. M.J. KLEIJN, *Demand differentiation in inventory systems.*
190. J.F.M. SWEEGERS, *Coordination, cooperation and institutions.*
191. P.A. DE HEK, *Economic growth and fluctuations: A theoretical enquiry.*
192. J. MEIJAARD, *Decision making in research and development: A comparative study of multinational companies in the Netherlands and the United States.*
193. K.P.B. OLDENKAMP, *Derivatives in portfolio management.*
194. C.M.M.P. WETZELS, *Squeezing birth into working life. Household panel data analyses comparing Germany, Great Brittain, Sweden and the Netherlands.*
195. J.M. VAN LEEUWEN, *Armoede en koopkrachtongelijkheid; theoretische beschouwingen over het effect van inkomensafhankelijke prijzen.*
196. P. FRIJTERS, *Explorations of welfare and well-being.*
197. J.A. CABRAL VIEIRA, *The evolution of wage structures in Portugal 1982-1992.*
198. M.J. SMITS, *Technology choice and the role of institutions. The case of wheat and the input market in Pakistan.*
199. M.T. KÖGEL, *Services and economic growth.*
200. D.J.C. VAN DIJK, *Smooth transition models: Extensions and outlier robust inference.*
201. E. KAPER, *Panel effects in consumer research - Statistical models for underreporting.*
202. K. SADIRAJ, *Albania: Transition to a market economy.*
203. A. RANASINGHE, *A pearl of great price: The free education system in Sri Lanka.*
204. P.J. VAN DER SLUIS, *Estimation and inference with the efficient method of moments: With applications to stochastic volatility models and option pricing.*
205. P.G.W. ALDERS, *Family ties in an aging world. Causes and consequences of fertility shifts.*
206. P. VAN HASSELT, *Dynamics of price formation in financial markets.*
207. K. VERWEIRE, *Performance consequences of financial conglomeration with an empirical analysis in Belgium and the Netherlands.*
208. P.W.T. GHIJSEN, *Labour and technology in Japan. An analysis of labour adjustment and technological change.*
209. E. DRISSEN, *Government decisions on income redistribution and public production. A political-economic general equilibrium approach.*
210. J. SPREEUW, *Heterogeneity of hazard rates in insurance.*