THE APPLICATION OF DATA ENVELOPMENT ANALYSIS TO CREDIT RISK MEASUREMENT

Adamandia Paraskevi Caldis

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Johannesburg, 2007

DEDICATION

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

____

____________ day of ____________

ABSTRACT

In competitive markets only the strong survive. For an institution to survive it needs to achieve high levels of performance through continued improvements and learning. Credit risk measurement has come under intense scrutiny with the recent Basel II Accord, an Accord that obliges banks to seek more efficient means for the management of their credit risk. In this dissertation I examine Data Envelopment Analysis (DEA), its extensions and applications to credit risk measurement. DEA is a performance measurement technique used to evaluate *relative efficiency* of a group. It is a quantitative model with a solid mathematical and economic underpinning; solving several linear programs simultaneously. Its greatest advantage is a feature that allows it to process multiple inputs and multiple outputs, thus uncovering relationships which remain hidden from other methodologies. We consider DEA as a useful tool for improving credit risk measurement within peer group analysis, a supplement or complement to the credit ratings and a validation tool for credit ratings. We apply DEA models to two credit risk measurement areas: Corporate Credit Risk and Country Risk. In the first application we apply DEA to corporate entities and compare efficiency to corporate credit ratings. Overall we find efficiency and credit rating have common elements. This confirms the common sense notion that corporates receiving better credit ratings are more efficient. In the second application we apply DEA to countries. We identify those countries that are performing better than the rest based on their efficiency. We discover that when a country's efficiency is compared to country credit rating, efficiency and credit rating are measuring two different but equally important aspects. We find that efficiency is a good indicator of relative economic acceleration. We observe that efficiency has the potential to identify future improvements in country credit quality. Based on the application of DEA to Corporate and Country Credit Risk, we recommend the use of DEA as a complementary tool to credit rating models and a means of facilitating better measurement and management of credit risk within banks.

DEDICATION

To my loving family Efstathios, Maria, Tina, Costa Caldis and my pillar Demetris Paphitis.

ACKNOWLEDGEMENTS

Special thanks are due to my supervisors Professor D. Sherwell and Mr C. Myburgh for constant guidance, support and valuable insight.

Contents

List of Figures

List of Tables

1

Introduction

1.1. Background

Banks are constantly looking for new methods and tools that can better manage their credit risk. Credit risk has become one of the central topics in risk management today. The last decade has seen substantial research on practical models to support an entity's creditworthiness assessment, pricing of high credit risk instruments, measurement and control of credit exposures and analysis of portfolio credit losses. Within South Africa, credit risk management has come under the spotlight with the new Basel II Accord [**17**].

The Basel II Accord is a set of regulations developed by the Basel Committee of banking supervisory authorities. The Basel Committee was established by the central bank governors of the G-10 countries in 1975. It was created to promote greater consistency in the way banks and banking regulators approach risk management across national borders. Basel II improves the capital framework's sensitivity to the risk of credit losses generally by requiring higher levels of capital for those borrowers thought to present higher levels of credit risk, and vice versa. An earlier Accord, Basel I, adopted in 1988, is now widely viewed as outmoded as it is risk insensitive and can easily be circumvented by regulatory arbitrage. The Basel II deliberations began in January 2001, driven largely by concern about the arbitrage issues that develop when regulatory capital requirements diverge from accurate economic capital calculations.

In July 2006 a comprehensive version was published including all Basel regulations up to this date [**2**]. Implementation of the Accord is expected by 2008 in many of the over 100 countries currently using the Basel I Accord. South Africa is one of these 100 countries and banks are scrambling to meet the 2008 deadline.

A South African newspaper reported that the South African national treasury published the Banks Amendment Bill on 23 February 2007, which aims to bring South African banking legislation in line with the Basel II requirements [**18**]. 'The amendments are aimed at strengthening the supervision of the banks and improving co-operation and information sharing between the various supervisors' [**18**]. The newspaper reports that banks will have to become smarter about the way they do business and look for methods which give capital relief. Although the big four South African banks are probably among the global leaders in complying with the Basel II requirements, they still have much to do to become fully compliant [**18**].

The final version of the Basel II Accord aims at ensuring that capital allocation is more risk sensitive. It clearly separates operational risk from credit risk, and quantifies both. The Basel II Accord also attempts to align economic and regulatory capital more closely to reduce the scope for regulatory arbitrage. The Basel II Accord has been structured in three pillars to achieve the objectives mentioned above:

(1) minimum capital requirements
(2) supervisory review
(3) market discipline

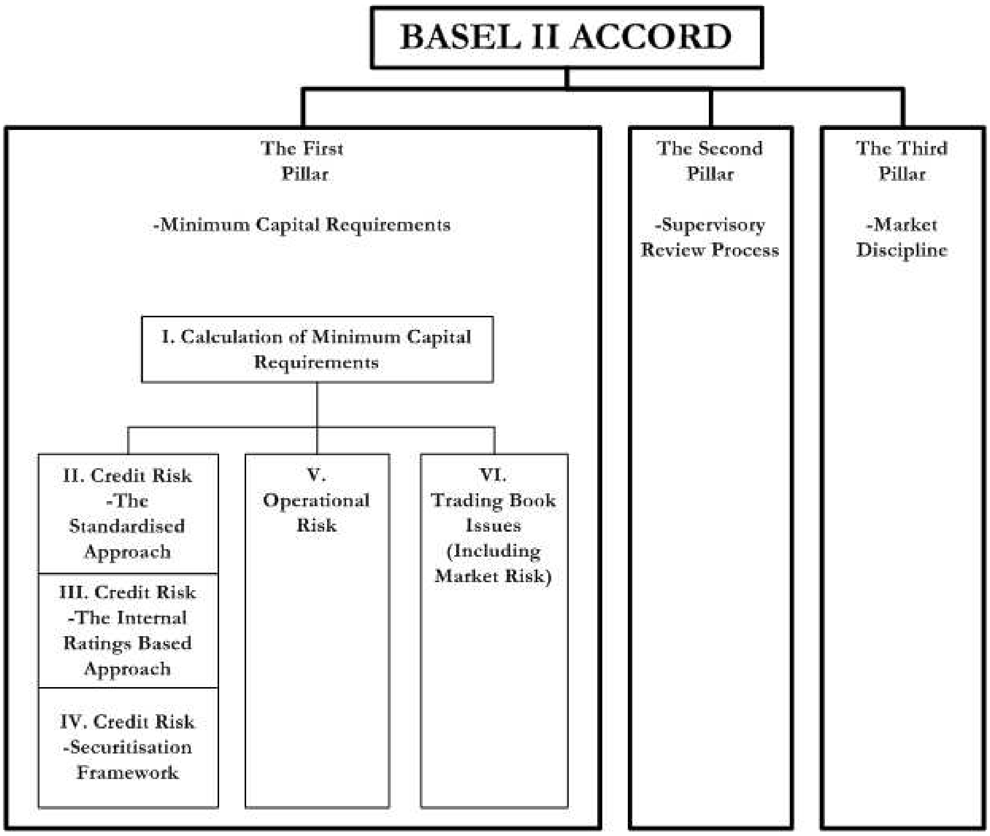

FIGURE 1.1. Basel II Pillars from the Basel II Accord [**2**]

The first pillar deals with improved risk sensitivity in the capital calculation requirements for three major risk components banks face: credit risk, operational risk and market risk. The Accord provides a selection of approaches (varying in sophistication) for calculating these risks. There are three approaches for calculating credit risk: The Standardised approach and the Internal Ratings Based (IRB) approach, where the later is further broken down into the Foundation Internal Ratings Based (FIRB) and Advanced Ratings Based (AIRB) approach. For the IRB approaches (the more sophisticated measures of credit risk), measures such as Expected Loss (EL) whose components are made up of Probability of Default (PD), Exposure at Default (EAD) and Loss Given Default (LGD) are to be calculated. PD as defined by the Basel II Accord is the greater of, the likelihood of an entity defaulting within one year or 0.03% for a bank or corporate exposure [**2**]. EAD simply is the total amount outstanding at the time of default. LGD is the expected loss in the case of default. A glossary of acronyms and banking terms is given at the end of this Dissertation. For the IRB approach PD must be calculated, EAD and LGD are prescribed. For AIRB all three risk components must be calculated. The calculation of these components requires advanced data collection and sophisticated risk management techniques.

The second pillar deals with the regulatory response to the first pillar, giving regulators much improved 'tools' over those available to them under Basel I. It also provides a framework for dealing with all the other risks that a bank faces, such as name risk, liquidity risk and legal risk, which the Accord combines under the title of residual risk.

The third pillar increases the disclosures that the bank must make. This pillar allows the market to have a better picture of the overall risk position of the bank and to allow entities to price and deal appropriately.

By far the largest portion of the Basel II Accord focuses on credit risk. It details the concepts of *internal (credit) risk ratings* produced by rating systems or rating models. Risk ratings are rating buckets used to differentiate between differing credit risk quality. The most commonly known credit risk ratings are given by rating agencies such as Fitch, Moody's and S&P. These agencies generally use lettered rating buckets from 'AAA' (excellent credit quality) to 'C' (poor credit quality) and default grade 'D'. Internal risk ratings are simply those ratings produced internally by a bank to facilitate it in its' credit risk management and are not disclosed to outsiders. Banks' internal ratings are somewhat like ratings produced by Fitch, Moody's, S&P, and other public rating agencies in that they summarise the risk of default or loss due to failure by a given borrower to pay as promised. However, banks' rating systems differ significantly from those of the agencies (and from each other) in architecture and operating design.

Most large banks use ratings in one or more key areas of risk management that involve credit, such as guiding the loan origination process, credit approval processes, portfolio monitoring and management reporting, analysis of the adequacy of loan loss reserves or capital, profitability, loan pricing analysis, and as inputs to formal portfolio risk management models. In short, internal risk ratings are the primary summary indicator of risk for banks' individual credit exposures. They both shape and reflect the nature of credit decisions that banks make daily. Having said this, Basel II, knowing the importance of this indicator, has provided detailed regulations on how these ratings must be derived, their design, operations, governance and formalised the role they must play in daily use.

Banks generally use internal or external credit rating models to produce risk ratings. These models are either statistically based, expert judgement (based on expert opinion) or a hybrid of the two. They can range from purely subjective to purely objective models, and generally combine quantitative and qualitative information to obtain a rating and PD. To give a tangible example, corporate rating models generally contain financial information with subjective qualitative adjustments such as management quality, where applicable behavioral factors are included. These are combined and modeled to produce a rating.

These models are commonly overlaid with additional features such as 'early warning signals', overrides (human intervention) and others. Early warning signals take into account those once-off events leading to default or downgrades. Overrides occur when human intervention overrides a model result with the intention to capture significant, entity-specific issues that are not covered by the model.

Credit ratings are produced by quantifying the credit risk exposures at either the entity or facility level. Each rating grade then maps to a specific PD range. For example a rating of 5 may be mapped to a PD of 2%. This would mean that a corporate rated as a 5 will have a 2% probability of default. We could also view PD in another way. If there are 100 corporate entities rated 5, two of the 100 corporates are expected to default. PD is a continuous measure. When this rating is mapped to a PD the rating can be viewed as an ordered categorical variable. This is important to highlight, as we will want to compare the rating to another measure later on.

When producing internal ratings, the Basel II requirements state that banks must put in place a 'rating system', which comprises all of the methods, processes, controls, data collection and IT systems that support the *assessment of credit risk, the assignment of internal risk ratings, and the quantification of default and loss estimates* i.e. PD, EAD, LGD estimates. Banks' rating systems and methodologies must have the ability to rank, order and quantify risk in a consistent, reliable and valid fashion [**2**]. This must be applied to all asset classes. Asset classes are a Basel II categorisation of exposures, namely corporate, sovereigns, banks, equities and retail. Corporates are further divided into large corporates, Small-to-Medium Enterprises (SME), Project Finance (PF), Object Finance (OF), Commodity Finance (CF), Income-Producing Real Estate (IPRE), High-Volatility Commercial Real Estate (HVCRE) [**2**].

Banks are required to associate their internal ratings to their own PD estimates (and own LGD estimates if the AIRB approach mentioned above is adopted). Details around the rating system design involve adhering to the rating dimensions (distinction of an entity dimension and a facility dimension); the rating structure, which involves the internal ratings having meaningful distributions (no excessive clustering or concentrations), a minimum of seven internal grades and one default grade for corporate, sovereigns and banks. For specialised lending a minimum of four internal grades and one default grade is required. All internal ratings must be associated to PD estimates which must be derived using the rating criteria. This involves the use of relevant information for rating and having a clear process in place. Where information is unavailable conservatism must be applied. The rating assignment horizon used by banks in PD estimation is expected to be longer than one year. Where credit rating models are the primary or partial basis of rating, sufficient human oversight is necessary. The rating models must undergo a regular cycle of model validation. Model validation involves the monitoring of model performance and a stability review of model relationships, that includes testing of model outputs against actual outcomes [**2**].

The rating system operations involve: rating each individual borrower and guarantor; assessing the integrity of rating process and independent assignment of ratings; review of ratings at least annually; the rules around human intervention, when a rating is produced (known as overrides); data maintenance and finally stress tests to identify future changes in economic conditions [**2**].

The rating system and the internal ratings are further subject to requirements regarding corporate governance and oversight i.e. approval of material aspects by board of directors, review by Internal or adequate External Audit, and risk quantification and validation [**2**].

The task of validating the rating system and the internal ratings is larger than that of the rating process. There is no one way to validate, but these methods must be consistent, well articulated standards [**2**].

> 'Banks must have a robust system in place to validate the accuracy and consistency of rating systems, processes, and the estimation of all relevant risk components. A bank must demonstrate to its supervisor that the internal validation process enables it to assess the performance of internal rating and risk estimation systems consistently and meaningfully.' [**2**] par 500.

The validation framework must test several items from backtesting to benchmarking and stability of the rating system. Backtesting is the comparison of estimated PD, EAD and LGD to actual realised and observed defaults, exposures and loss results.

Benchmarking can be described as taking a measurement against a reference point. For example in a ratings context it can be the comparison of a bank's internal credit rating to an agency rating given to an individual corporate. It is the search for best practices to improve an organization's products and processes. One can use benchmarking to assess relative performance, establish organisational targets and goals, monitor and learn from industry best practices. [**16**]. Where data is scarce or in the case where portfolios experience a low number of defaults, backtesting results become inappropriate. In these situations benchmarking and benchmarking tools become increasingly important. South African banks are particularly focused on using benchmarking techniques as they are largely exposed to emerging markets were data is scarce and at times unreliable.

By far the most challenging feature that banks face in the validation requirements is that the validation methods must not vary systematically with the economic cycle. This critical one liner lies innocently in the Accord waiting to be overlooked.

> 'Banks must demonstrate that quantitative testing methods and other validation methods do not vary systematically with the economic cycle.' [**2**] par 503.

As industry cycles occur in the economy, rating models are also cyclical by construction. The level of cyclicality varies significantly between rating models. A cyclical model does not always predict the cycle correctly. It is this that validation must test. Another reason Basel II prescribes that validation methods must not be cyclical, is to ensure that at any instant in the economic cycle and everything else being equal, the banks validation method produces the same result.

We have highlighted the importance of banks finding new tools to meet the Basel II requirements. Now that we have an understanding of Basel II and concepts such as credit risk management, credit ratings, credit rating models, PD, validation and benchmarking, we are in the position to discuss the objective of this dissertation.

1.2. Problem and Approach

The objective of this dissertation is to investigate and apply Data Envelopment Analysis (DEA) to credit risk measurement. DEA will facilitate banks who are constantly looking for new methods and tools that can help them meet their Basel II requirements and better manage their credit risk. DEA is a non-parametric, performance measurement technique used for evaluating *relative efficiency* with advantages over commonly used parametric approaches such as regression analysis. DEA is a tool with potential to facilitate banks to better manage their credit risk in several ways: in peer group analysis; as an input into credit rating models; as a supplementary tool to credit rating models, such as 'early warning signals' which can be used to indicate expected corporate credit failure or 'business opportunity signals' which can be used to identify where large profits can be made or indicate where to focus the business in the future and finally as a validation or benchmarking tool for credit ratings that meet the Basel II regulation of not varying systematically with the economic cycle. DEA models have several advantages: they are objective, quantitative and make use of well established economic principals.

We will apply DEA models to two credit risk measurement areas. The first application will be within the South African bank corporate lending sector. This is relevant due to the growing volume, variety and complexity of corporate businesses. Banks exposed to this sector are constantly looking for methods and tools which facilitate their current models that judge corporate performance and likely default. In a South African context very little default data exists or has been captured. In the second application we will apply DEA models to country credit risk. Part of assessing corporate behavior, performance and default also depends on the country's performance and likely default. This application will help us understand the country risk.

We will calculate efficiency results for corporates and countries and compare them to credit ratings. The comparison of DEA results to credit ratings will raise several questions and will supply information about how best to use the DEA results, for example, in peer group analysis, as an input into credit rating models, as a supplementary tool to credit rating models, as a validation or benchmarking tool.

The outlined approach to this dissertation will begin with an introductory chapter to the basics of DEA, its mathematical formulation, strengths and limitations. In Chapter 3 we will gain a greater understanding of DEA and its recent developments and extensions. Once we have fully understood DEA, we will apply it to the corporate lending sector in Chapter 4 and analyse the results. In Chapter 5 we will follow a similar application only this time to countries. Chapter 6 will conclude our findings and highlight the best uses of DEA in credit risk measurement.

2

Data Envelopment Analysis

2.1. Overview

Data Envelopment Analysis (DEA) is a performance measurement technique which is used for evaluating the *relative efficiency* of individual decision making units (DMUs) by using multiple inputs and multiple outputs. DMUs is a term introduced to cover, in a flexible manner, a collection of entities utilising similar inputs to produce similar outputs to be evaluated by a DEA model [**8**]. DEA applies linear programming methods to observed data to locate the best-practice frontier and construct a *non-parametric* piece-wise surface over the data [**4**]. This surface is also known as the efficient frontier or envelopment surface. The envelopment surface represents the revealed best-practice production frontier. Efficiency measures are then calculated for each DMU relative to this envelopment surface. To be efficient the DMUs must lie on the envelopment surface. DMUs that do not lie on the envelopment surface are termed inefficient. For an inefficient DMU, DEA identifies the sources and level of inefficiency for each of the inputs and outputs i.e. the sources of inefficiency can be analysed and quantified for every evaluated unit.

DEA was originally developed by Charnes, Cooper and Rhodes (1978) to create a performance measure that managers could use when conventional market-based performance indicators were unavailable [**5**]. DEA was used to evaluate productive efficiency among non-profit entities and then spread to evaluate the relative productive efficiency of branches in large networks and of individual institutions in entire industries [**16**]. More recently uses of DEA have involved evaluating performances of many different kinds of entities engaged in many different activities in many different contexts in many different countries [**8**]. The range of entities span from business firms, governments to non-profit agencies including courts, criminal justice systems, hospitals, schools, countries, regions, police forces and military units [**8**]. Recent publications such as [**8**] and [**12**] go to great lengths to consolidate current developments and continued applications of DEA.

DEA is a methodology directed to frontiers rather than central tendencies. The focus of DEA is on the individual observations in contrast to the focus on averages and estimation of parameters associated with regression approaches. Because of this feature, DEA has proven to be useful in uncovering relationships that remain hidden from other methodologies [**4**].

> 'DEA is an alternative and a complement to traditional central-tendency (statistical regression) analyses, and it provides a new approach to traditional cost-benefit analyses and frontier (or best-practices) estimation. DEA is a linear programming-based technique that converts multiple inputs and multiple outputs into a scalar measure of relative productive efficiency.' [**16**]

Efficiency as originally used in engineering and science represents the usual output-to-input ratio. For example in the field of combustion engineering where efficiency is the ratio for the actual amount of heat liberated in a given device to the maximum amount which could be liberated by fuel used is $E_r = \frac{y_r}{y_R}$ where y_R is the maximum heat that can be obtained from a given input of fuel, and y_r is the heat obtained from this same fuel input by the device being rated [**5**]. In this case y_r is considered as an output and y_R an input. The result obtained would yield a value of E_r between zero and unity. In common with most engineering definitions we have confined our development to ratios of single outputs and single inputs, or predetermined weighted sums thereof [**5**]. DEA expands the ratio definition of efficiency to accommodate multiple outputs and multiple inputs. This is accomplished through not predefining weights but rather the choice of weights is determined directly from observational data subject only to a set of constraints [**8**]. In this way DEA measures *relative efficiency* by identifying the 'best-performers', creating the efficient frontier and comparing a DMU with all DMUs.

The efficiency measure of each DMU is a score between zero and unity and is defined as the weighted sum of outputs divided by the weighted sum of inputs [**19**]. A DMU with a score of one is termed efficient. DMUs with scores between zero and unity represent the 'degree of efficiency'. The weights produced by DEA have beneficial uses, for instance we can identify the changes in the level of input and output needed to achieve efficiency. They can also be used to perform 'sensitivity analysis' to determine changes across all DMU's data that will result in reclassification of DMUs from 'efficient' to 'inefficient'. DEA 'sensitivity analysis' can also be used to identify the ranges of data variation that can be allowed before a reclassification occurs [**8**], [**20**].

DEA detects *Pareto efficient* DMUs [**5**]. Pareto efficiency or Pareto-Koopmans efficiency means that a DMU is efficient if and only if it is not possible to improve any input or output without worsening some other input or output [**8**]. The origins of Pareto-Koopmans efficiency comes from the definition of Pareto optimality which formed the basis of welfare economics. Pareto optimality is defined as "a position in an economy from which it is impossible to improve anyone's welfare by altering production or exchange without impairing someone else's welfare" [**8**]. Pareto optimality later extended to include 'production economics' that includes efficiency prices (related to our weights discussed above) and then further extended in DEA to what we earlier described as Pareto-Koopmans efficiency.

We now have two definitions of efficiency, the engineering-science efficiency and the Pareto-Koopmans efficiency. Although we may initially assume these definitions are quite different they are actually mathematically related to each other. We mentioned that DEA utilises linear programming and we know we can solve linear programs using either the primal or the dual formulation. Engineering-science efficiency and Pareto-Koopmans efficiency are mathematically related through the *duality theorem* of linear programming [**8**].

Now that we have defined efficiency there is a further differentiation we need to discuss within the DEA context. DEA does not measure economic efficiency but rather technological or productive efficiency. Productive efficiency is involved with the measure of levels of inputs relative to levels of outputs - therefore only input and output data is required. For example for an entity to be productively efficient it must minimise its inputs given outputs or either maximise its outputs given inputs. Economic efficiency encompasses but is not limited to productive efficiency as it involves optimally choosing the levels and mixes of inputs and outputs based on reactions to market prices. Therefore it requires input and output data as well as market price data. An entity would optimise an economic goal such as cost minimisation or profit maximisation, if it seeks to be economically efficient. According to [**16**] it is possible that some productively efficient entities are economically inefficient and vice versa. They believe this depends on the relationship between manager's abilities to use the best technology and their abilities to respond to market signals. This leads them to believe that economic efficiency requires both allocative and productive efficiency. For our purposes DEA will be used to measure relative productive efficiency although DEA can be adapted to examine economic efficiency [**3**],[**10**].

2.2. Mathematical Formulation

Figure 2.1 illustrates the basic concept of DEA graphically, using a very simple one input one output DEA model for nine DMUs. Each DMU utilises different amounts of input to produce various amounts of output. DEA produces *relative* efficiency measures for each DMU, these are plotted in Figure 2.1 for each of the nine DMUs. The efficiency measures are *relative* because DEA generates them from actual observed data for each DMU [**4**], [**20**]. The efficiency measures of DMUs A, B, E and H lie on the solid line known as the efficient frontier and are termed efficient. DMUs C, D, F, G and I are inefficient, since their input and output coordinates are located on the interior of the efficient frontier. In this simple example it is clearly visible that the efficient frontier *envelops* the data - we can see how the name data envelopment arises.

Let us use a geometric argument to explain why DMU A, B, E and H are efficient. Consider efficient DMU B. When we compare DMU B to DMU C which lies vertically below B, we see that B and C have the same amount of input, but B produces more output. This simple fact makes DMU B more efficient than C. When we compare DMU B to DMU F which lies horizontally to its right, we see that B produces the same amount of output as F, but with less input. This makes B more efficient than F. If a DMU existed to the left of DMU B, that produced the same output with even less input, then B would be dominated by this new DMU and no longer form the efficient frontier. Since this DMU does not exist and DMU B is non-dominated, by pareto dominance, B is efficient and dominates all

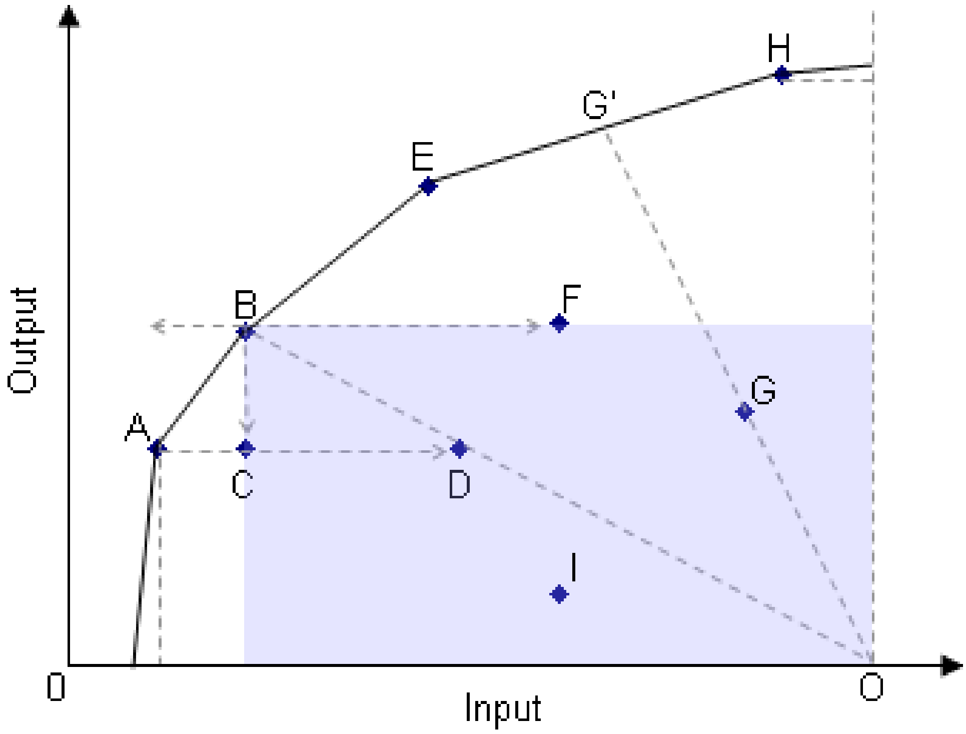

FIGURE 2.1. A graphical illustration of DEA for a one input one output model.

DMUs in the shaded region as seen in Figure 2.1 i.e. DMUs that produce equal or less output with equal or more input because B achieves more output with the least amount of input. Now consider DMU A. DMU A does not lie in the shaded region of DMU B. We find we cannot compare DMU A and B because although A produces less output than B it also uses less input. We observe that DMU A is also non-dominated. There are no DMUs that produce more output with as little input as used by A, so DMU A is efficient and forms part of the efficient frontier. For the same reasons demonstrated above DMU E and H are also efficient because each is non-dominated in the sense that their input and output co-ordinates lie on the convex hull of the set of data points.

Earlier we mentioned that the level of efficiency is measured relative to the efficient frontier. Therefore for each inefficient DMU, it is possible for DEA to identify the sources and amount of inefficiency in terms of input and output and provide a summary measure of relative productive efficiency [**16**]. Point G' is a projection of DMU G onto the efficient frontier. G' is known as the target where E and H are G's efficient peers. The relative efficiency of DMUs C, D, F, G and I are measured radially from O. For example G is measured radially as OG/OG'. In measuring the efficiency as OG/OG' in DEA the DMU's input and output geometrically determines the weights used in measuring it's efficiency. We have just demonstrated graphically how DEA achieves a direct connection between mathematics and economics, whereby the mathematical results have economic interpretations. Later we will discuss how other mathematical techniques lack this critical feature.

Now that we have geometrically observed that DEA has a solid economic and mathematical underpinning let us delve into the theory. The theory of DEA is centered around the 'classical' efficiency tests [**5**] which have the formulation below [**6**]. Note that the Euclidean inner product notation is used for x and y in $\mathbb{R}^n$: $(x, y) = x^T y$.

Suppose that each of N DMUs operate with s inputs and r outputs. Consider the i-th DMU. Let $X^i \in \mathbb{R}^s$ denote the tuple of inputs and $Y^i \in \mathbb{R}^r$ denote the ordered set of outputs of the i-th DMU. Each DMU is tested for efficiency over the same common set, i.e. for $k = 1, 2, ..., N$ we solve the following mathematical program:

$$(P_k) \begin{cases} \displaystyle\max_{u^k \in \mathbb{R}^r, v^k \in \mathbb{R}^s} \frac{(u^k, Y^k)}{(v^k, X^k)} \\ s.t. \\ \displaystyle\frac{(u^k, Y^{k'})}{(v^k, X^{k'})} \leq 1, \qquad\qquad k' = 1, 2, \ldots, N. \\ u^k \geq 0, \quad v^k \geq 0. \end{cases} \tag{2.2.1}$$

Mathematically, the relative productive efficiency of each DMU is calculated as the ratio of its total weighted output over its total weighted input. DEA selects the weights that maximise each DMU's productive efficiency measure. In this process each DMU maximises its efficiency by choosing its individual weights which suit it best to its situation. When doing this it ensures that two set of constraints are met. The first constraint is that no ratio, weighted sum of outputs over weighted sum of inputs, can be greater than one. The notation of this constraint is key. It is telling us that by design the set of weights obtained to maximise each DMU must also be feasible for all the other DMUs included in the calculation, hence the use of the notation k for the weights (u and v) and k' for the outputs (Y) and inputs (X) of other DMUs. Any entity is able to utilise the same set of weights to evaluate its own efficiency ratio. It is also important that every other DMU has an efficiency less than one with the chosen weights of DMU i because the DEA measure should be scaled between zero and one [**16**]. Geometrically we can see in Figure 2.1 that this constraint imposes a specific normalization or boundary around the data. It has no other significance, since we can see that the results of the maximization problem are independent of any particular choice of units. The second constraint ensures that all weights are non-negative. Since inputs, outputs and their weights are positive we notice that our program is formulated in the positive quadrant.

To compare a number of DMUs among a variety of criteria simultaneously, the same number of separate, but related, mathematical programming problems must be optimised and the results so combined [**16**], this is represented as (P_k).

The formulation of (P_k) above is a fractional programing problem with the disadvantage that it can not be solved by a linear program. The fractional problem can be transformed into an equivalent ordinary linear programming problem. We derive the ordinary linear program with a transformation that multiplies by the denominator: $\xi^k = (v^k, X^k)^{-1}$, $w^k = \xi^k v^k$, $z^k = \xi^k u^k$. We can linearise because by construction ξ^k is positive. Therefore,

$$
\text{(2.2.2)} \qquad \begin{cases} \displaystyle\max_{\frac{z^k}{\xi^k} \in \mathbb{R}^r, \frac{w^k}{\xi^k} \in \mathbb{R}^s} \xi^k \left(\frac{z^k}{\xi^k}, Y^k\right) \\ s.t. \\ \qquad \xi^k = \dfrac{1}{(\frac{w^k}{\xi^k}, X^k)} \\ \qquad \dfrac{(\frac{z^k}{\xi^k}, Y^{k'})}{(\frac{w^k}{\xi^k}, X^{k'})} \leq 1, \qquad\qquad k' = 1, 2, \ldots, N. \\ \qquad \dfrac{z^k}{\xi^k} \geq 0, \quad \dfrac{w^k}{\xi^k} \geq 0. \end{cases}
$$

Notice the additional constraint introduced by the transformation in program (2.2.2). By isolating the common factors we obtain

$$
\text{(2.2.3)} \qquad \begin{cases} \displaystyle\max_{z^k \in \mathbb{R}^r, w^k \in \mathbb{R}^s} \frac{\xi^k}{\xi^k}(z^k, Y^k) \\ s.t. \\ \qquad \dfrac{\xi^k}{\xi^k}(w^k, X^k) = 1 \\ \qquad \dfrac{1}{\xi^k}(z^k, Y^{k'}) \leq \dfrac{1}{\xi^k}(w^k, X^{k'}), \qquad k' = 1, 2, \ldots, N. \\ \qquad \dfrac{z^k}{\xi^k} \geq 0, \quad \dfrac{w^k}{\xi^k} \geq 0. \end{cases}
$$

Note that the objective function is now linear in z and the above N tests are simplified and reduce to the N linear programs:

$$
\text{(2.2.4)} \qquad (P'_k) \begin{cases} \displaystyle\max_{z^k \in \mathbb{R}^r, w^k \in \mathbb{R}^s} (z^k, Y^k) \\ s.t. \\ \qquad (w^k, X^k) = 1 \\ \qquad (z^k, Y^{k'}) \leq (w^k, X^{k'}), \qquad k' = 1, 2, \ldots, N. \\ \qquad z^k \geq 0, \quad w^k \geq 0. \end{cases}
$$

The k-th DMU is said to be efficient if, and only if, the optimal value of the program (P'_k) is equal to one. Geometrically this occurs when the DMU lies on the efficient frontier as seen in Figure 2.1.

Some authors replace the inequalities $z^k \geq 0$, $w^k \geq 0$ by $z^k \geq \epsilon e$, $w^k \geq \epsilon e$, where e is the vector of ones and $\epsilon > 0$ is a non-Archimedean quantity [**6**], [**12**]. Geometrically this means that the efficient frontier can not be vertical or horizontal on either edge. Notice in Figure 2.1 that the efficient frontier from DMU A downwards is not vertical and from DMU H to the right is not horizontal. This is due to the restrictions that ensure strictly positive i.e. non-negative and *non-zero* weights.

Figure 2.1 is shown to help visualise DEA in two dimensions for a simple one input, one output DEA model. In general and as seen by our mathematical formulation, (P'_k), we will deal with data with higher dimensionality (i.e. beyond three dimensions) and cannot visualise the efficient frontier graphically, therefore our results are presented in tables.

2.3. Illustrative Example of DEA

For purposes of explaining the DEA approach, an illustrative example of its application is given below. Assume we wish to identify which corporate entities operating in the agricultural industry are efficient. Financial information of a set of nine corporate entities (agricultural corporate A to I) was sourced. Based on industry experience, we select key drivers that we believe influence agricultural corporate entities efficiency. Let these be the following financial information: Total Assets, Total Debt, Total Sales and Net Income. Of these four, we can argue that the first two are corporate business inputs, as the corporate business is based on the assets it owns and debt that is a method of how it chooses to fund its business. We could also argue that the last two indicators are corporate outputs, as the debt and assets that a corporate has chosen to fund its business has generated total sales and net income. A DEA model is built based on nine linear programs and the total result of outputs over inputs is maximised. Two inputs and two outputs for each of the nine agricultural corporates is fed into the DEA model as specified in Table 2.1 and the DEA weights and efficiency measures in Table 2.2 and Table 2.3 are produced.

In Table 2.3 corporates A, B and D are efficient. If this example could be visualised in four dimensional space, then A, B and D would form the efficient frontier similar to what we observed in Figure 2.1. It is important re-emphasize that DEA only gives the relative efficiencies i.e. efficiencies relative to the data considered. It does not, and cannot, give us absolute efficiencies or economic efficiency for the reasons discussed above.

Corporate ID	Inputs (X^i)		Outputs (Y^i)	
	Total Assets ('000)	Total Debt ('000)	Total Sales ('000)	Net Income ('000)
A	1,838,077	201,562	4,053,142	262,403
B	90,141.00	38,188	275,124	3,773
C	3,887,810	466,474	6,246,885	149,516
D	2,273,832	1,915	4,026,998	232,434
E	766,356	94,130	988,610	33,616
F	190,093	33,122	359,737	15,974
G	163,742	10,410	179,658	11,003
H	300,842	116,114	311,690	3,901
I	179,385	42,048	138,245	4,927

TABLE 2.1. An illustrative example of identified input and output fields to be used in a DEA model.

Corporate ID	Input Weights (w^i)		Output Weights (z^i)	
A	ϵ	3.811	0.544	ϵ
B	3.635	ϵ	11.094	ϵ
C	0.115	ϵ	0.220	0.311
D	ϵ	4.302	ϵ	522.190
E	0.582	ϵ	1.112	1.571
F	2.211	ϵ	4.221	5.965
G	3.021	ϵ	5.341	12.053
H	1.127	ϵ	2.151	3.039
I	2.193	ϵ	4.188	5.917

TABLE 2.2. An illustrative example of the input and output weights determined by the DEA model.

Corporate ID	Efficiency
A	1
B	1
C	0.720
D	1
E	0.576
F	0.795
G	0.543
H	0.351
I	0.303

TABLE 2.3. An illustrative example of efficiency measures determined by the DEA model.

2.4. Strengths and Uses

To demonstrate the simplicity of using DEA and how objective its results are, let us compare it to two techniques commonly used and widely known: ratio analysis and regression analysis.

A simple and commonly used method for comparing companies is ratio analysis. Typically an output measure is divided by an input measure. Using the example above consider Total Debt divided by Total Assets more commonly known as the Debt Ratio. The Debt Ratio is a financial leverage measure that indicates what proportion of debt a company has relative to its assets. A Debt Ratio greater than one, indicates that a company has more debt than assets. A Debt Ratio less than one indicates a company has more assets than debt. Debt Ratios provide information about protection of creditors from insolvency and the ability of firms to obtain additional financing for potentially attractive investment opportunities. The more debt a company has, the more likely it is that the company will become unable to fulfill its contractual obligations (too much debt can lead to a higher probability of insolvency and financial distress). Let us compare the same group of corporates in the agricultural industry using the Debt Ratio.

Corporate ID	Efficiency	Debt Ratio	Net Return on Assets
A	1	0.110	0.143
B	1	0.424	0.042
C	0.720	0.120	0.038
D	1	0.001	0.102
E	0.576	0.123	0.044
F	0.795	0.174	0.084
G	0.543	0.064	0.067
H	0.351	0.386	0.013
I	0.303	0.234	0.027

TABLE 2.4. Ratio analysis using Debt Ratio and Net Return on Assets compared to DEA

The Debt Ratio reveals that all the corporates have a Debt Ratio less than one in Table 2.4. Corporate B appears to have the highest Debt Ratio of 0.424, and corporate D the lowest Debt Ratio of 0.001. Does this mean that Corporate B is more likely to become unable to fulfill its contractual obligations and eventually lead to financial distress? The Debt Ratio does not distinguish between the corporates very well and so we must investigate another ratio. A single ratio does not give a clear picture and is only one view of a company. Suppose we have two views of a company, Debt Ratio and a profitability ratio such as Net Return on Assets. Net Return on Assets is calculated by dividing Net Income by Total Assets. The higher the Net Return on Assets the more profitable the company. Table 2.4 reveals that corporate A has the highest Net Return on Assets and H the worst Net Return on Assets. In the Debt Ratio E has a lower Debt Ratio than F and therefore E outperforms F. In the Net Return on Assets Ratio, F has a higher Net Return on assets than E and therefore F outperforms E. If we were to rank corporates E and

F, would E rank higher than F or visa versa? We have just experienced a common problem when using ratio analysis. Adding another ratio may help understand more about the individual companies but will certainly create more confusion. Different ratios give a different picture and it is difficult to firstly select the relevant ratios and secondly to combine the entire set of ratios into a single numerical judgment. Should they weight equally? Is leverage perhaps more important than profitability in our example? Further, when the same sample is given to a group of individuals to analyse independently, consistency is lost because of subjectivity. One individual may deem one of the ratios more important than another individual assessing the companies and therefore come to a different conclusion.

Although ratio analysis is still widely used within many South African banks, the process is generally inadequate, as simple ratios are used and human judgement grossly compromises consistency when combining these ratios. DEA on the other hand can handle selected multiple inputs and multiple outputs to produce an efficiency score. DEA objectively selects and combines the weightings for the inputs and output measures by maximising the efficiency, unlike the subjective combination of ratios within ratio analysis.

Another method that is widely used in credit rating models is a statistical technique called regression analysis. This is a *parametric approach* which utilises all the information available in the data to determine dependent parameter(s) based on a number of independent parameters. Its objective is to *fit a single regression hyperplane* through the data. This single optimised regression equation is assumed to apply to all DMUs. Regression analysis is a typical statistical approach, that is characterised as a central tendency approach which evaluates producers relative to an average producer. For example it will create a *'mythical' average* company that scores average in all the financial ratios and compare every other company relative to this mythical average company.

The first limitation we experience with applying this technique is that it requires the imposition of a specific functional form relating the independent variable(s) to the dependent variable. For the example above, we would need to make a choice between a logistic regression equation or a probit regression[**4**]. With either choice of functional form selected, they require specific assumptions about the distribution of the error terms in particular independently and identically normally distributed. From our example above, the financial data collected has its own distribution and imposing an assumption of normal distribution may not be reflective of the 'real' situation. Assuming we have solved the first limitation we encounter another. There can only be a single output (dependent variable), therefore where there is more than one output we must combine them. Combining all outputs into a single indicator function can be problematic. How do we combine net income and total sales as one measure? They are measuring two different things. There are ways to combine these two measures, but what would this combination mean. As in any combination, information is lost, assuming we have managed to combine all the output factors into a meaningful single output like a rating or efficiency for example. Although the results may supply the efficiency or rank order, they do not lend themselves to any other meaningful interpretations, such as what can be done to the inputs and outputs to improve inefficient DMUs i.e. it can not make connections with economics (Pareto solutions). Regressions also suffer from scaling effects. A small co-efficient may be interpreted as a factor being insignificant but this may be due to unscaled units. For example if our regression involved using Total Assets and the Debt Ratio. Total Assets has a range between R90 million to R183 million, where as Debt has a range between 0.001 and 0.424. The regression may apply a small co-efficient to the Debt Ratio making this factor insignificant. By far the largest limitation is that it measures efficiency relative to *'mythical' average* performances. For DEA, the DMUs are directly compared against a peer or combination of peers. We can compare Regression Analysis and DEA graphically. This is illustrated in Figure 2.2. The DEA efficient frontier lies on a piecewise surface on top of the observations on actual efficient DMUs. Now observe that statistical regression analysis tries to fit a regression line through the centre of the data. The regression line creates a mythical average performance in the sense that no DMUs exist on this regression line yet all are compared to it.

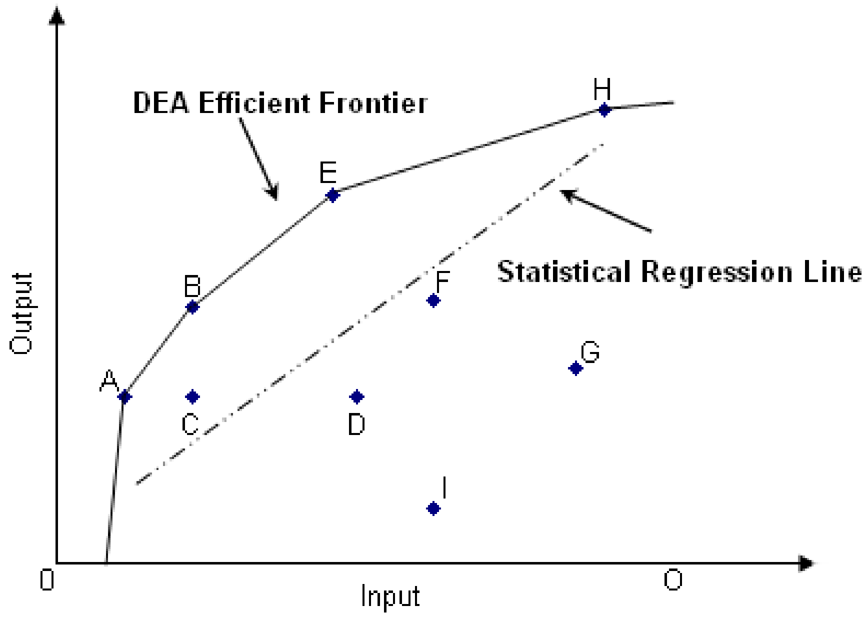

FIGURE 2.2. Comparison of DEA and Regression Approaches [**16**]

Essentially regression analysis is a 'global' efficiency comparison. It tries to fit all entities to one equation. It is applicable for a sample with fairly homogeneous entities where their data meets the regressions assumptions and no further information about improvement of the entity is required.

Unlike regression analysis that is a *parametric approach*, DEA does not require the arbitrary selection of a functional form relating to inputs and outputs, or any other assumptions [**4**]. DEA is a *non-parametric* form of estimation; that is, no *a priori* assumption on the analytical form of the production function or distribution assumptions are required. DEA can handle multiple input and multiple output models. This avoids calculating a single measure of input or output. As seen in the example above both Total Sales and Net Income were outputs. Inputs and outputs can have very different units. For example variable x_1 could be in units of lives saved and x_2 could be in units of rands without requiring a priori trade-off between the two. DEA can determine the amount of input to be used or the magnitude of output to be achieved for each entity to become fully efficient. It compares all DMUs to the efficient DMUs unlike regression analysis that compares every DMU to a 'mythical' average DMU.

We have seen the benefits of DEA over ratio analysis and regression analysis. Let us discuss some other benefits and uses of DEA.

DEA can also be used as a benchmarking tool that provides much more information [**19**]. Generally, as benchmarking analysts we are left to our own devices as to how to analyse data, characterise and measure gaps and project future performance levels. The objective of any benchmarking exercise is to provide information for improvements. So we are now at the mercy of the technique used for benchmarking and totally reliant on the amount of information the technique provides. Commonly used benchmarking techniques provide simple gap analysis. Although these give us important insights they also miss many. Single gap analysis examines only one measure at a time and ignores any interactions between key variables. For example they provide a one-dimensional view of service, product and processes but are totally unable to capture interactions between these key variables. What we need to capture this relationship is an innovative approach that has a multiple input, multiple output benchmarking framework. Data Envelopment analysis is exactly this type of approach. It is a multiple input, multiple output technique and because of this and the strengths mentioned above provides more information than a simple benchmark. [**16**] states that DEA can be useful when adapted for benchmarking purposes. It provides an analytical, quantitative benchmarking tool for measuring relative productive efficiency. It has proven to be a valuable tool for strategic, policy and operational decision problems, particularly in the service and non-profit sectors [**16**].

In summary, simple one dimensional gap analysis is not sufficient. A better benchmarking paradigm should have the following characteristics: a solid economic and mathematical underpinning; alternative actual and composite/hypothetical best practice units; the ability to take into account the trade-offs and substitutions among the benchmark metrics and lastly a means to suggest directions for improvement on the many organisational dimensions included in the study.

This leads us to another use of DEA. If DEA is used as a benchmarking technique then it automatically qualifies to be used in validation. Where Basel II has set out clear validation requirements, DEA can be used as it meets the criteria of being an objective, quantitative tool that does not vary systematically with the economic cycle.

DEA may also prove useful in peer group analysis. It is best practice for banks to assess credit risk not only on an individual basis but also on an industry level, known as peer group analysis. Industries undergo cyclical effects, the objective of peer group analysis is to eliminate gain or loss made by 'all' good times or bad times respectively. Within an industry it aims to determine which entities are performing and which are not irrespective of the cycle the industry is in. DEA naturally lends itself to this application because it differentiates between efficient and inefficient entities within an industry, irrespective of the economic cycle.

Depending on the type of input and output information the DEA model receives it can also be used as an 'early warning signal' overlaying the core credit rating models. DEA produces efficiency results. Where the entities it is evaluating prove to be inefficient we could flag these entities as having the potential to deteriorate or default based on their inefficient operations. DEA could also prove to be an ideal tool for spotting efficient entities and potentially highlighting high rewards if we invest in these entities early. This could be used as an input or supplementary tool to the credit rating models.

In summary of its strengths, DEA can be thought of as a 'local' efficiency comparison, as it compares entities with the structurally most comparable peer entities and (in small samples) entities with exceptional characteristics may be shown as efficient 'by default'. The case that leads most to its advantage is that it allows multiple input and output analysis. It is ideal when applied to large samples with diverse entities. DEA overcomes the shortcomings of traditional financial ratio analysis, regression analysis and simple one dimensional gap analysis used in benchmarking. DEA can be used as an alternative to these methods. DEA can also be used in peer group analysis, validation and an input or supplementary tool into credit rating models.

2.5. Limitations

When used unwisely the same characteristics that make DEA a powerful approach can also create problems. DEA only gives us relative and technical efficiencies of DMUs. DEA converges very slowly to absolute efficiency. In other words referring to the illustrative example above, Corporate A, B and D are doing well compared to the rest of their peers in the group, corporates C, E, F, G, H and I, but this is not compared to any 'theoretical maximum'.

Since DEA is a non-parametric approach, statistical hypothesis tests are difficult and are the focus of on-going research. We might ask what type of hypothesis tests should we use. Take the corporate efficiencies in the illustrative example above for A to I. We may want to compare them to another measure such as the credit risk ratings assigned to each corporate. A typical hypothesis test would look to prove or disprove whether there is a relationship between efficiency scores and credit ratings. Table 2.5 shows the DEA efficiency scores and the credit ratings. In this example a rating of 1 implies excellent credit quality and a rating of 21 implies very poor credit quality or highly likely to default. In Table 2.5, it appears that the less efficient corporates have higher credit ratings i.e. they are more risky. Unfortunately we can not use a hypothesis tests to prove this.

Corporate ID	Efficiency	Credit Rating
A	1	12
B	1	15
C	0.720	14
D	1	10
E	0.576	13
F	0.795	13
G	0.543	15
H	0.351	17
I	0.303	16

Table 2.5. DEA efficiency scores compared to corporate credit ratings

Notice that there is one corporate that appears to be an outlier in this example. Corporate B is efficient according to DEA but has a poor credit rating of 15. When investigating a little further, all the corporates listed under Agriculture included crop and cattle farming, fishing and forestry, but corporate B was more involved in the production of animal products not necessarily farming. Therefore corporate B is not truly a corporate in the agricultural industry but rather in the food and beverages industry. This example highlights a limitation of DEA. Outliers may appear efficient. By combining DEA with a benchmark as done above, we can identify outliers, thus turning this limitation into a strength.

2.6. A Relevant Literature Study

This literature study sparked the idea for this dissertation. The study is interested in benchmarking the productive efficiency of U.S banks. It initially discusses the limitations of simple one-dimensional benchmarking methods and then identifies that DEA is an innovative approach, which overcomes these limitations. By using DEA as its benchmarking tool it concludes that:

> 'More efficient banks tend to be higher performers and safer institutions' [**16**].

It achieves this by comparing the volume of services provided and resources used by each bank with those of all other banks. It also uses the DEA results to compare them to CAMEL ratings. CAMEL (Capital adequacy, Asset quality, Management quality, Earnings ability and Liquidity) ratings are a common measure used by regulators for evaluating a bank's financial health [**20**]. The findings of the comparisons are of interest to us.

As any business in a competitive market will know, to be successful we do not only need to be continually improving and learning, we need to know our competitor. We need to compare ourselves to them and understand where we stand relative to them and the rest of the industry. Benchmarking measures can provide us with some of this information but we are at the mercy of the benchmarking method we choose to use. The most commonly used simple gap analysis (a one dimensional benchmarking method) highlights our strengths and weaknesses in key variables such as service, product and processes but misses interactions, substitutions and trade-offs between them [**16**]. So they give us half the story. What is needed to complete the story is a multiple-input, multiple-output benchmarking framework that can determine technical efficiency and provide information to inform business decisions on how to become well-managed. DEA is this type of tool with these characteristics and [**16**] uses a constrained-multiplier, input orientated DEA model to achieve a robust quantitative foundation to benchmark the efficiency of U.S banks.

The study focuses on the U.S commercial banking industry. They conducted a survey with twelve experienced Federal Reserve Bank of Dallas bank examiners and constructed a five input three output DEA model. This model selected variables that capture the financial intermediation functions of a bank with the objective of attracting deposits and making loans and investments. The inputs are defined as resources required to operate a bank. They are:

(1) Salary Expense
(2) Premises and Fixed Assets,
(3) Other Non-Interest Expense,
(4) Interest Expense,
(5) Purchased Funds.

The three outputs are defined as desired outcomes. They are:

(1) Earning Assets,
(2) Interest Income,
(3) Non-Interest Income.

Upper and lower bounds for the unit weights were also selected. 'According to this model, productively efficient banks, or best-practice banks, allocate resources and control internal processes by effectively managing their employees, facilities, expenses and sources and uses of funds while working to maximise earnings assets and income.'[**16**]. This model was applied to 11,397 banks in operation for the year end 1991, 10,224 banks in 1994 and 8,628 banks for 1997.

The study divided the data into quartiles based on efficiency scores to evaluate the input and output factors driving the efficiency results. They divided the input and output factors by Total Assets and this revealed several observations. In 1991 the most efficient banks had significantly lower Salary Expense, Premises and Fixed Assets, other Non-Interest Expense, and Purchased Funds, and they had significantly higher relative levels of Earning Assets, Interest Income, and Interest Expense. By 1997 only Interest Expense, Premises and Fixed Assets and Purchased Funds had statistically significant differences among the input variables. From 1991 to 1997, a significant advantage was found for the most efficient banks for Non-Interest Income. The most efficient banks earned a significantly higher Return on Average Assets, held higher Equity Capital and managed relatively smaller loan portfolios that tended to have fewer risky assets than the least efficient institutions. When the data was divided into asset size quartiles, they found significant differences in 1991 between large and small banks but not in 1994 or 1997.

The DEA results were then compared to confidential bank examiner CAMEL ratings. CAMEL ratings are an evaluation of a bank's financial health on a scale of one to five. A CAMEL rating of 1 is defined as an institution that is basically sound in every respect and a CAMEL rating of 5 is defined as an institution with critical financial weaknesses that render the probability of failure extremely high in the near term. The factors used to assess a banks CAMEL ratings are

- Capital adequacy
- Asset quality
- Management quality
- Earnings ability
- Liquidity

The banks were divided into two groups: Strong banks with a CAMEL rating of 1 or 2 and weak banks with a CAMEL rating of 3, 4 or 5. [**16**] found that the strong banks had significantly higher efficiency scores when compared to the weak banks.

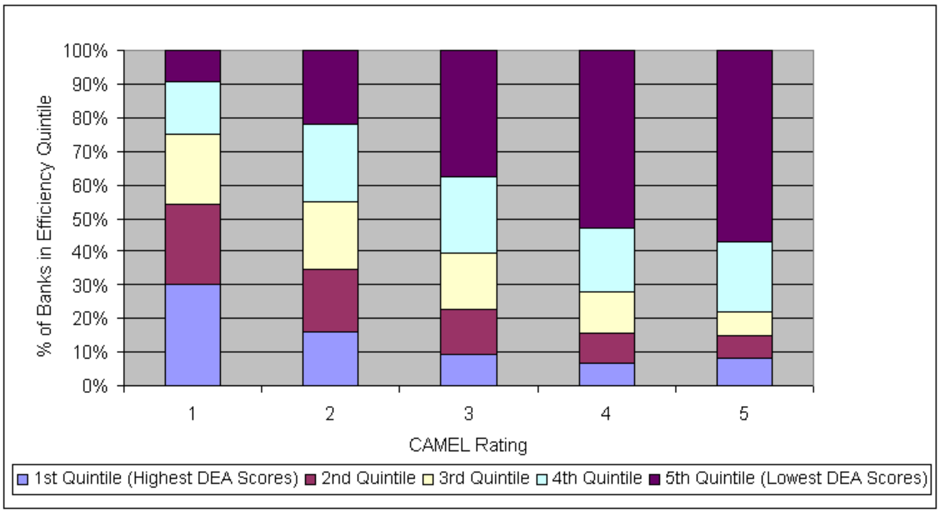

FIGURE 2.3. Efficiency Score Quintiles by CAMEL Rating Combined Data 1991, 1994, 1997 from [**16**].

In Figure 2.3, they streamed the banks into efficiency quintiles and diagrammatically show the percentage of efficient banks in each quintile within each CAMEL rating. If the DEA results did not differentiate between the banks in each CAMEL rating then we would expect each CAMEL rating to contain 20% of each DEA quintile. Instead what we see is the most efficient quintiles are overrepresented in the CAMEL rating 1 and the least efficient quintile overrepresented in the CAMEL rating 5. From this [**16**] deduces that because of the close association of CAMEL ratings and efficiency results the DEA model may be useful as an additional off-site surveillance tool for bank regulators.

This study concludes that DEA is a superior benchmarking tool when compared to common gap analysis and can be used as an indicator to benchmark performance. It shows that banks with better CAMEL ratings are more efficient. Because of this relationship, the DEA efficiency model can be used as an additional tool by regulators as well as bank managers to better understand their bank's productive abilities relative to the industry and its competitors.

> 'We have shown that more efficient banks tend to be higher performers and safer institutions' [**16**].

In this dissertation we will be trying something very similar with South African corporate entities to unveil a similar relationship if it exists. We will build a DEA model and compare the corporate credit ratings assigned by credit experts to DEA efficiency scores. There are several questions we will be looking to answer. Can DEA be used as a supplement, validation or alternative to credit ratings? Can we use it to benchmark against credit ratings? Can it be a supporting tool that explains more about the corporate market and the competition? Is DEA completely independent to credit ratings and should it be incorporated into the rating or even the business process. Ultimately, will it draw similar conclusions to [**16**]? Are more efficient corporates better performers and safer institutions?

3

DEA Extensions

3.1. Cross-Efficiency

As previously discussed, efficiency is simply a ratio of the weighted sum of outputs over the weighted sum of inputs. DEA selects the best set of weights for each DMU in order to maximise it efficiency score based on a set of constraints. These constraints ensure that no weight is negative and the resulting ratio is not greater than one. The efficiencies we have just described we will term 'simple efficiencies' and denote as $E_{k,k}$. 'Simple efficiency' is a special case of cross-efficiency.

Assume our DEA solution has selected the best set of weights for a particular DMU. Now use this same set of weights for all the other DMUs in the group. The result is a cross-efficiency of each of the other DMUs, as viewed by our original DMU. Calculating the cross-efficiencies for each DMU row-by-row in this manner will result in a matrix of cross-efficiencies as seen in the Table 3.1 below [**9**]. In Table 3.1 the leading diagonal is our original set of 'simple efficiencies'. $E_{3,2}$ is the efficiency of DMU 2 measured using DMU 3's weights.

DMU	DMU receiving weights						Average appraisal
	1	2	3	4	5	6	of peers
1	$E_{1,1}$	$E_{1,2}$	E_{13}	$E_{1,4}$	$E_{1,5}$	$E_{1,6}$	A_1
2	$E_{2,1}$	$E_{2,2}$	$E_{2,3}$	$E_{2,4}$	$E_{2,5}$	$E_{2,6}$	A_2
3	$E_{3,1}$	$E_{3,2}$	$E_{3,3}$	E_{34}	$E_{3,5}$	$E_{3,6}$	A_3
4	$E_{4,1}$	$E_{4,2}$	$E_{4,3}$	$E_{4,4}$	$E_{4,5}$	$E_{4,6}$	A_4
5	$E_{5,1}$	$E_{5,2}$	$E_{5,3}$	$E_{5,4}$	$E_{5,5}$	$E_{5,6}$	A_5
6	$E_{6,1}$	$E_{6,2}$	$E_{6,3}$	$E_{6,4}$	$E_{6,5}$	$E_{6,6}$	A_6
	e_1	e_2	e_3	e_4	e_5	e_6	
	Average appraisal by peers						

TABLE 3.1. DEA Cross-Efficiency matrix as displayed in [**9**]

If we average down the columns and exclude the DMUs 'simple efficiency' i.e. how each DMU rates itself with its best selected weights, we will be calculating an averaged peer appraisal by each DMU. In Table 3.1 this is represented as e_k. We will call e_k the peer efficiency. If we average along the rows we will be calculating the average appraisal of how each DMU views its peers. We will represent this as A_k. Note that the leading diagonal is not included in e_k or A_k averages. We will adopt the same intuitive notions as used by [**9**], where 'simple efficiency' is thought of as a process of self-appraisal and the calculated 'peer efficiency' above is thought of as peer appraisal of DMU k. What we need to make clear here is that our understanding of peer, is to any other DMU.

The reason we introduce the notion of peer efficiency is because it can act as a complement or alternative to the simple efficiency. Peer efficiency can be used to distinguish between the efficient DMUs. This is useful where a large portion of the population of DMUs are efficient. Referring back to our agricultural example, assume we would like a ranking of all our agricultural corporates. This is easy to do for the inefficient corporates, but for the three efficient corporate A, B and D we can use peer efficiency. See the ranking in Table 3.2. Notice that with the peer efficiency we have a ranking for our efficient DMUs A, B, and D. Further our peer efficiency has preserved the order for our inefficient DMUs.

Now that we have discussed simple efficiency and peer efficiency we can not continue without discussing mavericks. Mavericks are those DMUs that enjoy the greatest relative increment when shifting from peer efficiency to simple efficiency. If a DMU is classified as a maverick what we are implying is that this DMU operates 'far' from the rest of the DMUs. For example a DMU would be classified as a maverick if its simple efficiency is very high and its peer appraisal is very low leading to a low peer efficiency and therefore a high maverick indicator. We use this maverick indicator to identify those DMUs that are operating very differently and those DMUs that are all round performers.

Corporate ID	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)
D	1	0.866	0.155
A	1	0.841	0.189
B	1	0.760	0.316
F	0.795	0.641	0.240
C	0.720	0.542	0.328
E	0.576	0.447	0.289
G	0.543	0.434	0.251
H	0.351	0.264	0.330
I	0.303	0.241	0.257

TABLE 3.2. Ranking corporates with peer efficiencies

3.1.1. Mathematical Formulation of Cross-Efficiency.

To define cross-efficiency mathematically, let DMU k choose its own weights u^k for its outputs and v^k for its inputs. The cross-efficiency of DMU k', using the weights of k is defined as

$$E_{k,k'} = \frac{(u^k, Y^{k'})}{(v^k, X^{k'})}, \quad k' = 1, 2, \ldots, N, \quad k = 1, 2, \ldots, N \tag{3.1.1}$$

subject to the constraints

$$u \geq 0, \quad v \geq 0 \tag{3.1.2}$$

$$E_{k,k'} \leq 1 \quad \text{for all } k' \tag{3.1.3}$$

Average peer efficiency is defined as:

$$e_k = \frac{1}{(N-1)} \sum_{k' \neq k} E_{k,k'} \tag{3.1.4}$$

Maverick index is calculated in the following way:

$$M_k = \frac{E_{k,k} - e_k}{e_k} \tag{3.1.5}$$

3.2. Aggressive and Benevolent Formulation

Our primary objective is to select the best weights that maximise DMU k's simple efficiency. What is missing from our discussion is how the other DMU weights are assigned. As stated by [**9**] cross-efficiency may yield an arbitrary formulation, in that the weights which maximise DMU k's simple efficiency, arbitrarily assign weights to other DMUs, based on which the linear program has found solutions. To avoid this arbitrary assignment of weights to other DMUs, [**9**] has suggested a secondary goal. This secondary goal gives us a choice of how weights are selected. We could choose a solution that keeps DMU k as efficient as possible while making all other DMUs as inefficient as possible. This is known as the 'aggressive formulation'. Or we could take the other extreme and choose a solution that keeps DMU k as efficient as possible while enhancing the cross-efficiencies of other DMUs as much as possible. This is known as the 'benevolent formulation'.

Using our same agricultural corporate example we can view the effects of the aggressive and benevolent formulation for cross-efficiency in Table 3.2 and 3.3. In Table 3.2, the aggressive formulation, each DMU k will choose a solution that keeps it as efficient as possible while making all other DMU's as inefficient as possible. In Table 3.3, the benevolent formulation, each DMU k will choose weights that maintain its own efficiency while enhancing the cross-efficiency of the other DMUs as much as possible. There are two important points to notice in these Tables. The first is that in both Tables the diagonal is the 'simple efficiency'. It has not changed and is the same in both the aggressive and benevolent formulation. Remember we expect this to be the case as our primary objective is always to maximise DMU k's efficiency. The second point is that the benevolent formulation's cross-efficiencies are either equal to or higher than the aggressive formulation's cross-efficiencies.

Aggressive formulation

There are several ways to formulate the secondary goal to produce the aggressive and benevolent formulations [**9**], [**1**]. The most obvious method for an aggressive formulation would be to minimise the sum of the cross-efficiency fractions averaged down each column, but this requires us to solve a non-linear program. Here we will

Cross-efficiencies	Corporate ID								
Corporate ID	A	B	C	D	E	F	G	H	I
A	1.0000	0.8825	0.7152	1.0000	0.5713	0.7684	0.5428	0.3132	0.2851
B	1.0000	1.0000	0.7196	0.9265	0.5757	0.7954	0.5273	0.3511	0.3032
C	1.0000	1.0000	0.7193	0.9262	0.5756	0.7953	0.5272	0.3511	0.3032
D	1.0000	0.8825	0.7152	1.0000	0.5713	0.7684	0.5428	0.3132	0.2851
E	1.0000	1.0000	0.7195	0.9264	0.5757	0.7954	0.5273	0.3511	0.3032
F	1.0000	1.0000	0.7196	0.9265	0.5757	0.7954	0.5273	0.3511	0.3032
G	1.0000	0.8827	0.7153	1.0000	0.5714	0.7684	0.5428	0.3133	0.2851
H	1.0000	1.0000	0.7196	0.9265	0.5757	0.7954	0.5273	0.3511	0.3032
I	1.0000	1.0000	0.7196	0.9265	0.5757	0.7954	0.5273	0.3511	0.3032

Table 3.3. Benevolent formulation

present one of the simpler methods which can be implemented and solved with a standard linear program [**9**].

Assume we have determined our simple efficiencies $E_{k,k}$. Let us minimise the sum of the weighted outputs (numerator of the cross-efficiency) minus the sum of the weighted inputs (denominator of the cross-efficiency) as our secondary goal. We will define the secondary goal as B_k where

$$\min_{u\in\mathbb{R}^r, v\in\mathbb{R}^s} B_k = \sum_{k'\neq k}(u^k, Y^{k'}) - \sum_{k'\neq k}(v^k, X^{k'}), \tag{3.2.1}$$

subject to

$$(v^k, X^k) = 1, \tag{3.2.2}$$

$$E_{k,k'} \leq 1 \quad k' \neq k, \tag{3.2.3}$$

$$(u^k, Y^k) - E_{k,k}(v^k, X^k) = 0, \tag{3.2.4}$$

$$u \geq 0, \quad v \geq 0. \tag{3.2.5}$$

The benevolent formulation can be obtained by maximising the secondary goal B_k.

3.3. Ranking Efficient DMUs

Earlier we discussed cross-efficiencies and introduced peer efficiency as an alternative to simple efficiency. Peer efficiency as described by [**9**] allowed us to rank DMUs when DEA cannot provide enough information to rank efficient DMUs, but since 1994 more sophisticated methods have been developed. In 2004 Liu and Peng [**11**] published a model called Common Weight Analysis (CWA) to rank efficient DMUs. Described briefly, CWA determines a common set of weights for all DMUs as opposed to peer efficiency where each DMU selects its own best suited weights. They then introduced the Super Common Weight Analysis (SCWA) and later in 2006, [**12**] presented a method that determines the most compromising set of indices' weights analysis (MWCA). In this section we will discuss these three models. Later in the dissertation we will build linear programs for CWA, SCWA and MCWA method and apply them to our efficient DMUs where we would like further differentiation. We will compare CWA, SCWA, MCWA and peer efficiency.

CWA determines a common set of weights for all DMUs. It assumes all efficient DMUs are equally weighted. It creates an implicit efficiency datum line set equal to one. Each DMU is assigned an implicit efficiency which is less than or equal to one. Then CWA minimisies the total gaps to the datum line. To formulate CWA mathematically, let E represent the efficient DMUs. Let the common output weights for the efficient DMUs be $U_1, U_2, \ldots, U_r$ and the common input weights be $V_1, V_2, \ldots, V_s$. If Δ_k^O represents the total gaps of the outputs for the k-th DMU to the datum line and Δ_k^I the total gaps of the inputs for the k-th DMU to the datum line, then the CWA linear program is

$$\min \sum_{k \in E} (\Delta_k^O + \Delta_k^I), \quad k = 1, 2, \ldots, N \tag{3.3.1}$$

subject to

$$\sum_{i=1}^{r} y_{i,k} U_i - \sum_{c=1}^{s} x_{c,k} V_c + \Delta_k^O + \Delta_k^I = 0, \tag{3.3.2}$$

$$\Delta_k^O \geq 0, \quad \Delta_k^I \geq 0, \tag{3.3.3}$$

$$U_i \geq \epsilon > 0, \quad i = 1, 2, \ldots, r, \tag{3.3.4}$$

$$V_c \geq \epsilon > 0, \quad c = 1, 2, \ldots, s. \tag{3.3.5}$$

where ϵ is a very small positive coefficient [**12**].

The SCWA model is similar to the CWA model except that the implicit efficiencies can be greater than or equal to one. It is formulated as

$$\min \sum_{k \in E} (\Delta_k^O + \Delta_k^I), \quad k = 1, 2, \ldots, N \tag{3.3.6}$$

subject to

$$\sum_{i=1}^{r} y_{i,k} U_i - \sum_{c=1}^{s} x_{c,k} V_c - \Delta_k^O - \Delta_k^I = 0, \tag{3.3.7}$$

$$\Delta_k^O \geq 0, \quad \Delta_k^I \geq 0, \tag{3.3.8}$$

$$U_i \geq \epsilon > 0, \quad i = 1, 2, \ldots, r, \tag{3.3.9}$$

$$V_c \geq \epsilon > 0, \quad c = 1, 2, \ldots, s. \tag{3.3.10}$$

where ϵ is a very small positive coefficient [**12**].

The MCWA model creates a datum line equal to one. It determines the most compromising set of indices' weights which then calculate the implicit efficiencies for the efficient DMUs. These are either above, below or equal to one. MCWA then behaves like the ordinary least-square method (OLS) of the residuals in statistical regression analysis, by minimise the total of the gaps to the datum line [**12**]. Figure 3.1 shows this diagrammatically [**12**]. A and B are efficient DMUs with implicit efficiencies on the graph. The diagonal line is the datum line with a slope equal to one.

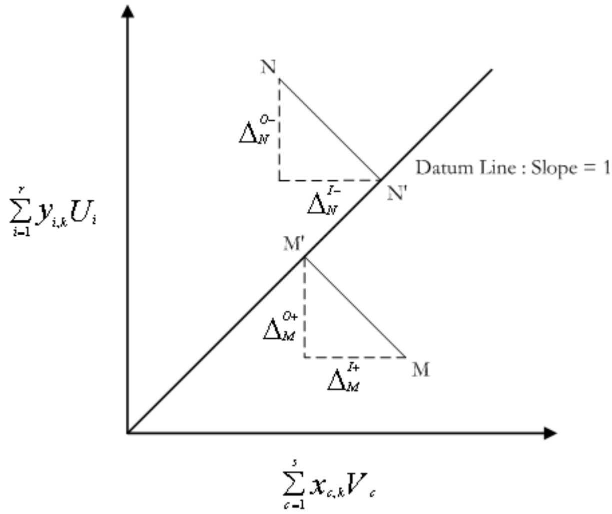

FIGURE 3.1. Gaps to the datum line in MCWA method from [**12**]

MCWA is formulated in a two phase approach. In the first phase we solve for Δ^*. Δ^* minimises the total of the gaps to the datum line.

$$\Delta^* = \min \sum_{k \in E} (\Delta_k^+ + \Delta_k^-), \quad k = 1, 2, \ldots, N \tag{3.3.11}$$

subject to

$$\sum_{i=1}^{r} y_{i,k} U_i - \sum_{c=1}^{s} x_{c,k} V_c + \Delta_k^+ - \Delta_k^- = 0, \tag{3.3.12}$$

$$\Delta_k^+ \geq 0, \quad \Delta_k^- \geq 0, \tag{3.3.13}$$

$$U_i \geq \epsilon > 0, \quad i = 1, 2, \ldots, r, \tag{3.3.14}$$

$$V_c \geq \epsilon > 0, \quad c = 1, 2, \ldots, s. \tag{3.3.15}$$

where the gaps of inputs and outputs are integrated into the gaps $\Delta_k^+ = \Delta_k^{O+} + \Delta_k^{I+}$, $\Delta_k^- = \Delta_k^{O-} + \Delta_k^{I-}$ and ϵ is a very small positive coefficient [**12**].

The first phase leads to the optimal value for Δ^* but in order to obtain a set of unique common weights we need to solve the second phase.

$$\min \sum_{i=1}^{r} U_i - \sum_{c=1}^{s} V_c \tag{3.3.16}$$

subject to

$$\sum_{i=1}^{r} y_{i,k} U_i - \sum_{c=1}^{s} x_{c,k} V_c + \Delta_k^+ - \Delta_k^- = 0, \tag{3.3.17}$$

$$\sum_{k \in E} (\Delta_k^+ + \Delta_k^-) = \Delta^*, \quad k = 1, 2, \dots, N \tag{3.3.18}$$

$$\Delta_k^+ \geq 0, \quad \Delta_k^- \geq 0, \tag{3.3.19}$$

$$U_i \geq \epsilon > 0, \quad i = 1, 2, \dots, r, \tag{3.3.20}$$

$$V_c \geq \epsilon > 0, \quad c = 1, 2, \dots, s. \tag{3.3.21}$$

where the gaps of inputs and outputs are integrated into the gaps $\Delta_k^+ = \Delta_k^{O+} + \Delta_k^{I+}$, $\Delta_k^- = \Delta_k^{O-} + \Delta_k^{I-}$ and ϵ is a very small positive coefficient [**12**].

Once we have solved for the weights in CWA, SCWA, and MCWA we can determine the ranking of the efficient DMUs using Θ_k^*. By sorting Θ_k^* in descending order we will obtain the most efficient to the least efficient DMU. Θ_k^* is calculated using the following equation:

$$\Theta_k^* = \frac{\sum_{i=1}^{r} U_i^* y_{i,k}}{\sum_{c=1}^{s} V_c^* x_{c,k}}, \quad k = 1, 2, \dots, N \tag{3.3.22}$$

When discriminating between the efficient DMUs we can select any model, CWA, SCWA and MCWA. We can not make any statements about which of the three methods is the best to use. We can briefly compare these three methods to peer efficiency. Peer efficiency has different properties, it is not only generated by the efficient DMUs but also the inefficient DMUs unlike CWA, SCWA and MCWA. The advantage that peer efficiency has over these three methods is that discrimination extends to the inefficient DMUs, but the disadvantage is that it contains an averaging effect which may not be desirable. In the next chapters we will calculate and demonstrate the use of CWA, SCWA and MCWA methods.

4

Application to Corporate Entities

In the previous chapters we have discussed the DEA methodology in detail. We are now in a position to apply DEA to corporate credit risk. The objective of this exercise is to determine if DEA can be used by banks to better manage their corporate credit risk, and if so, how should it be used. Can DEA be used as a supplement, complement, or alternative to credit ratings or in peer group analysis? Can it facilitate in explaining the corporate market and identify the competition? Are more efficient corporates better performers and safer institutions? In this chapter we will apply DEA to corporate entities and obtain efficiency scores as well as corporate rankings. We will analyse the efficiency scores and compare them to credit ratings where this is relevant. In this application we will investigate the information that DEA provides about our corporate entities and the relationships and differences between efficiency and credit ratings. We will also investigate if DEA can provide us with a method to rank corporate entities when credit ratings can not differentiate between them.

4.1. Sample Data

4.1.1. Data Collecting and Cleaning. The corporate data we collected was obtained from a large South African bank on signing a confidentiality agreement. The corporate data contains information about a large variety of South African corporate entities. These include corporates listed on the Johannesburg Securities Exchange (JSE) which we will term 'Listed corporates' and corporates which are not listed on the JSE which we will term 'Unlisted corporates'. Both Listed and Unlisted corporates are well represented in the sample data. There are two groups of Unlisted corporates: 'Unlisted-A' corporates which contain larger unlisted corporate businesses and 'Unlisted-B' corporates which contain smaller unlisted corporate businesses.

An industry classification has been assigned to each corporate by the credit managers. The industry classification indicates the corporate's core business. Where a corporate operates in several industries, the industry with the greatest source of income was assigned.

We have obtained the credit ratings for each corporate. The credit ratings are derived through a customised credit rating model that considers both quantitative and qualitative components. The credit ratings selected for this application exclude the influence of South African country risk and are standalone ratings. This means that the influence of other entities (Corporate or Government) support in the form of a guarantee or other implicit support is not taken into account. The credit ratings are ordered categorical numbers ranging from 1 to 21. A credit rating of 1 means excellent credit quality, a credit rating of 21 means very poor credit quality where the corporate entity is highly likely to default. Investment grade (highly rated) corporates receive credit ratings in the range of 1 to 10. The credit managers assigned the internal credit ratings to each corporate. The internal credit ratings represent the bank's view of the corporate's credit quality. This credit rating maps directly to a corresponding median PD for that credit rating's PD range. Although there is a mapping to the PD, it is not derived directly from a default data model and then divided into credit ratings. It is derived by ranking the corporates then dividing them into credit ratings and using default data to map the credit ratings to a median PD for each rating. What we are highlighting here is that the granularity of the ranking can only go as far as the number of credit ratings. The PD does not provide further granularity because of the way it was derived. Therefore common problems banks face where lack of default data has forced this type of derivation, is differentiating between corporates when they all cluster within one credit rating. On this issue we will look to DEA for potential solutions to differentiate between the corporates.

We obtained the latest available financial information for each corporate containing Balance sheet items such as: Total Assets, Cash and Near Cash, Tangible Assets, Intangible Assets, Total Liabilities, Total Debt, Long Term Debt, Total Equity, Ordinary Share Capital, Share Capital, Retained Earnings. Income statement items such as EBITDA, EBIT, Total Sales, Interest Expense, Net Income are obtained. One cash flow item is obtained, namely Funds from Operations (FFO).

Corporates were excluded from the data set if the financial information dated earlier than 2003 or if a large portion of the financial information was missing or undisclosed. After data cleaning, the final data set contained 190 corporates with complete financial information.

4.1.2. Sample Data Analysis. Analysing the 190 corporates along class, industry, rating and size reveals some interesting information. The Total Assets of these corporates range from between R9 million to R90 billion. Although the number of Listed corporates only make up 24% of the sample data as seen in Figure 4.1, they dominate Unlisted-A and Unlisted-B corporates by 63% in terms of sum of Total Assets as seen in Figure 4.2. Figure 4.2 also clearly shows the size difference between Unlisted-A and Unlisted-B corporates. Unlisted-B has slightly more corporates than Unlisted-A, but in terms of sum of Total Assets they are much smaller corporate businesses and only make up 1% of the sum of Total Assets.

Let us shift our attention to the corporate industries in Figure 4.3. The industries with the largest number of corporates are Retail General, Mining and Steel & Metals. The industries with the largest corporate sum of Total Assets are Telecommunications, Mining and Chemicals. Notice that the Retail General industry has far more corporates than any other industry, but looking at the sum of its corporate's Total Assets, it is only the forth largest. In the Retail General industry we

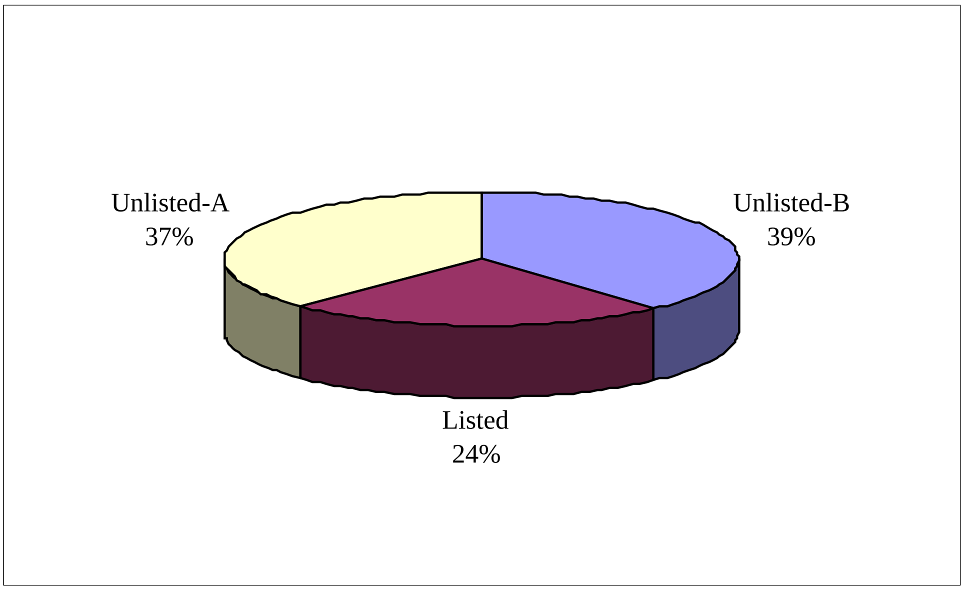

FIGURE 4.1. Number of Corporates in each Class (%)

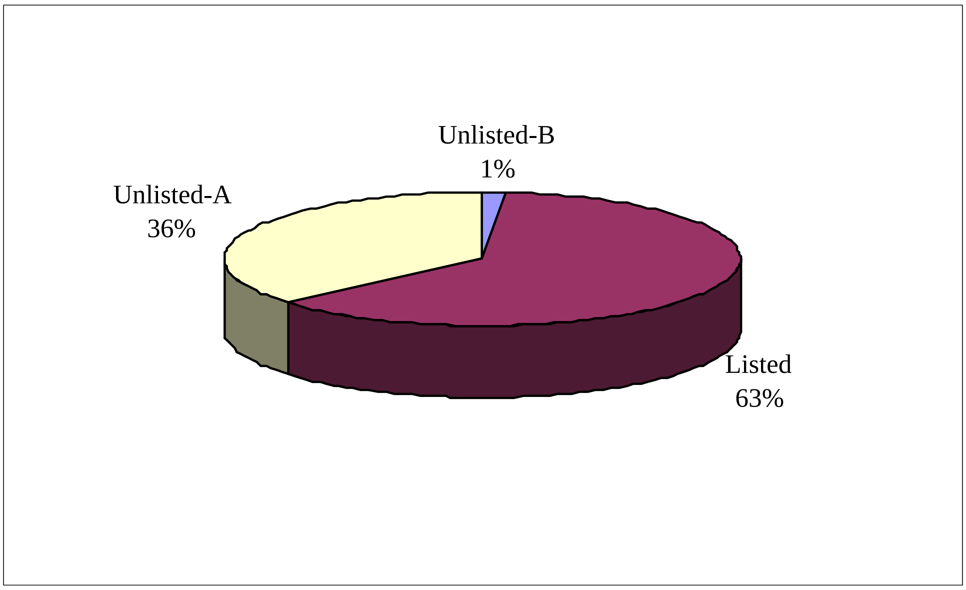

FIGURE 4.2. Sum of Total Assets of Corporates in each Class (%)

have many corporates with a small asset size. On the other hand, the Telecommunications industry is an industry with one of the smallest number of corporates but has the largest Total Assets. Telecommunications has few corporates but they have a large Total Assets size.

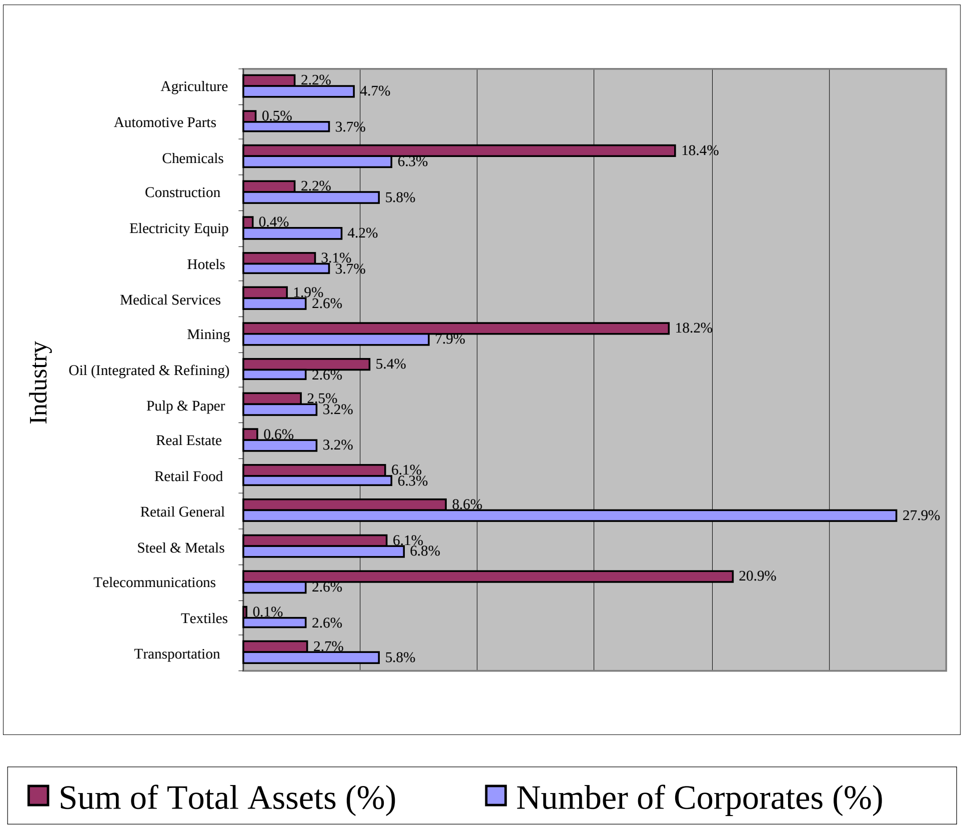

FIGURE 4.3. Corporate Industries

Analysing the credit ratings gives us an idea of the corporate credit quality. In Figure 4.4 we see that a large number of corporates are clustered around credit ratings 14 and 15, that is 45% of the corporates are rated either 14 or 15. This creates some real challenges if anyone lending credit would like to differentiate between the corporates within credit rating 14 or 15. Ignoring the under-populated credit ratings 13 and 16 (that was a consequence of a once-off mapping exercise), this graph resembles a normal distribution. Most credit ratings are produced using regression analysis models which assume normality assumptions, so it is interesting but not surprising that we see credit rating output of this nature, where most credit ratings are clustered within the centre of the credit rating scale. This coincides with the 'mythical' average constructed by regression analysis. Figure 4.5 has a very different shape, almost lognormal (again ignoring the under populated credit ratings 13 and 16). The best credit ratings 7 to 11 have a larger sum of Total Assets. Observing this fact in combination with Figure 4.4 where there are a few corporates in credit ratings 7 to 11, we can deduce that in this corporate sample data the largest corporates generally receive the best credit ratings. Investigating this relationship even further let us examine the number of corporate credit ratings within each class as seen in Figure 4.6. The Listed corporates which we saw earlier have the greatest sum of Total Assets and generally occupy the better end of the credit rating scale, then the Unlisted-A corporates follow. The Unlisted-B corporates with the smallest sum of Total Assets occupy the worst end of the rating scale.

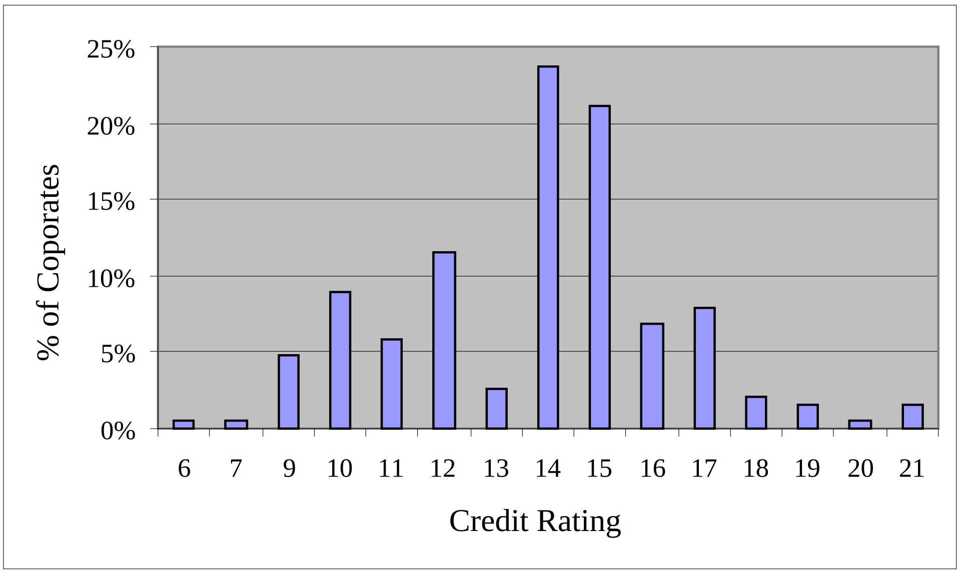

FIGURE 4.4. Number of Corporates in each Credit Rating (%)

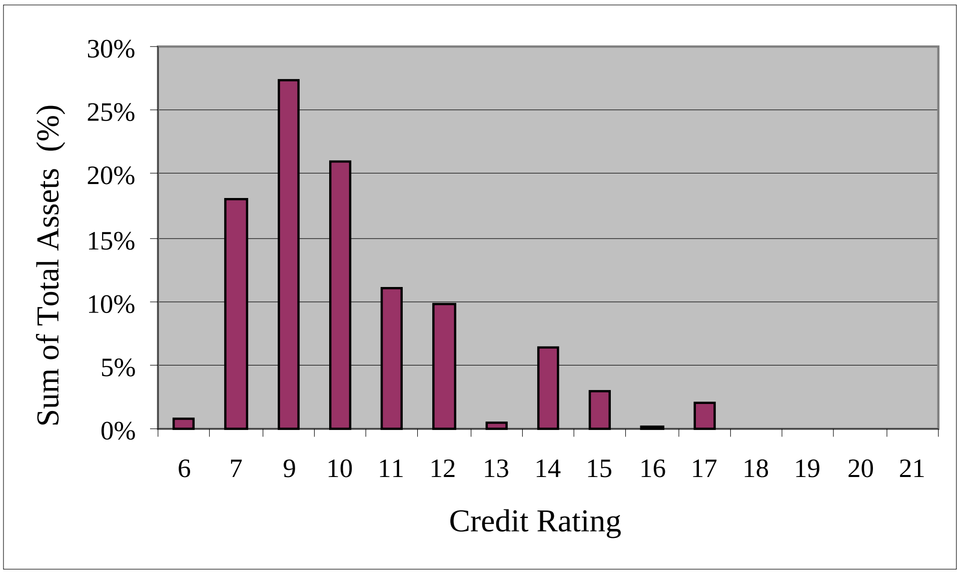

FIGURE 4.5. Sum of Total Assets in each Credit Rating (%)

Table 4.1 shows the distribution of ratings within each industry. Notice that different industries occupy different ranges on the credit rating scale. For example corporates in the Mining industry generally have better credit ratings than the

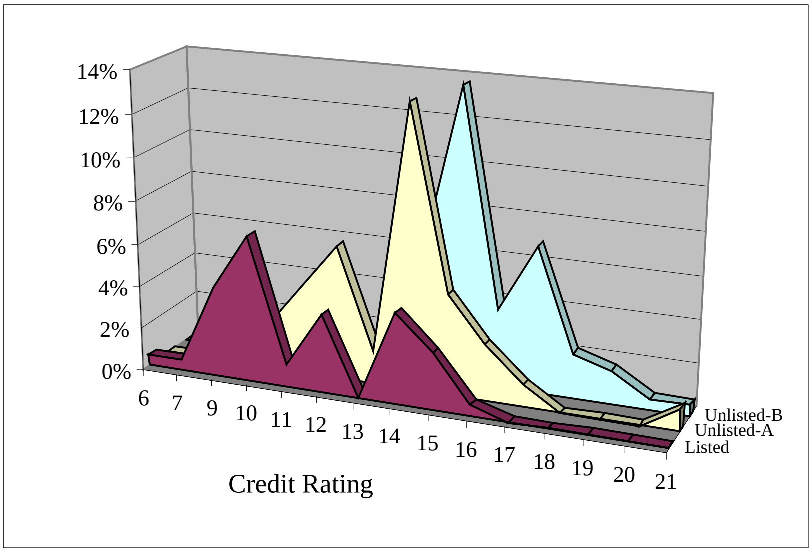

FIGURE 4.6. Credit Ratings distribution for each Class

Industry/Rating	6	7	9	10	11	12	13	14	15	16	17	18	19	20	21	
Agriculture				1	1	1	1	2	2	1						9
Automotive Parts			1			1		3	2							7
Chemicals		1					1	4	5		1					12
Construction						4		1	2	2	2					11
Electrical Equip								3	2	1	1		1			8
Hotels			1	2				3	1							7
Medical Services	1			1			1	1			1					5
Mining			3	3	2	1		2	3						1	15
Oil (Integrated & Refining)				1	2	1		1								5
Pulp & Paper						2		3			1					6
Real Estate				1				2			1		1		1	6
Retail Food			1	1		1		5	2	1	1					12
Retail General				4	4	7	2	9	13	4	5	4			1	53
Steel & Metals				1	1	2		2	5	1	1					13
Telecommunications			3		1			1								5
Textiles								1		3			1			5
Transportation				2		2		2	3		1			1		11
Grand Total	1	1	9	17	11	22	5	45	40	13	15	4	3	1	3	190

TABLE 4.1. Credit Ratings per Industry

Electrical Equipment industry except for one mine which is very close to default. Every industry has at least one corporate in rating 14 and the Retail General industry has most of it's corporate ratings clustered in and around credit rating 14 and 15. Transportation fills a wide range of the rating scale, this could be due to the diversity within this industry. For example Transportation includes air, land and sea so the corporates within this industry operate very differently and are exposed to different risks.

If we sort the corporates by size using Total Assets and divide them into four equally sized groups TA1 to TA4, where TA1 has the corporates with the largest Total Assets and TA4 has corporates with the smallest. By bucketing the credit ratings into four rating buckets as in Table 4.2, we can construct Figure 4.7.

Credit Rating	Rating bucket	Description
1 2 3 4 5 6 7 8 9 10	1	Investment grade corporates, very strong and least likely to default
11 12 13 14	2	Non-investment grade or strong corporates
15 16 17 18	3	Average corporates
19 20 21	4	Weak corporates, most likely to default

TABLE 4.2. Credit Ratings rolled up into Rating Buckets

This Figure shows the percentage of corporates within each rating bucket that falls into each Total Asset size bucket i.e. 46.8% of the largest corporates have a rating that places them in the best rating bucket 1. 46.8% of the largest corporates have the second best ratings and 6.4% the third best rating. Looking at the second largest Total Asset size bucket of corporates, TA2, we see that 10.6% have the best ratings, 80.9% have the second best ratings and 8.5% have the third best ratings. What we are seeing in TA1, TA2 and TA3 is a close association between rating and size. Based on this representative sample, it appears that a 'large' corporate is more likely have a good rating. Notice that this does not mean that if a corporate has a good rating, it is large, nor does it mean that if a corporate is 'small' it has a poor rating. This can be seen in TA4. In TA4 the corporates are 'small' with a total asset size that ranges between R9 million to R62 million, as apposed to TA3, TA2 and TA1 with size ranges from R62 million to R233 million, R233 million to R1.6 billion and R1.6 billion to R88 billion. TA4 having the corporates with the smallest Total Asset size also has ratings in all rating buckets from the best to the worst. In conclusion, if the corporate is 'large' between R62 million to R88 billion

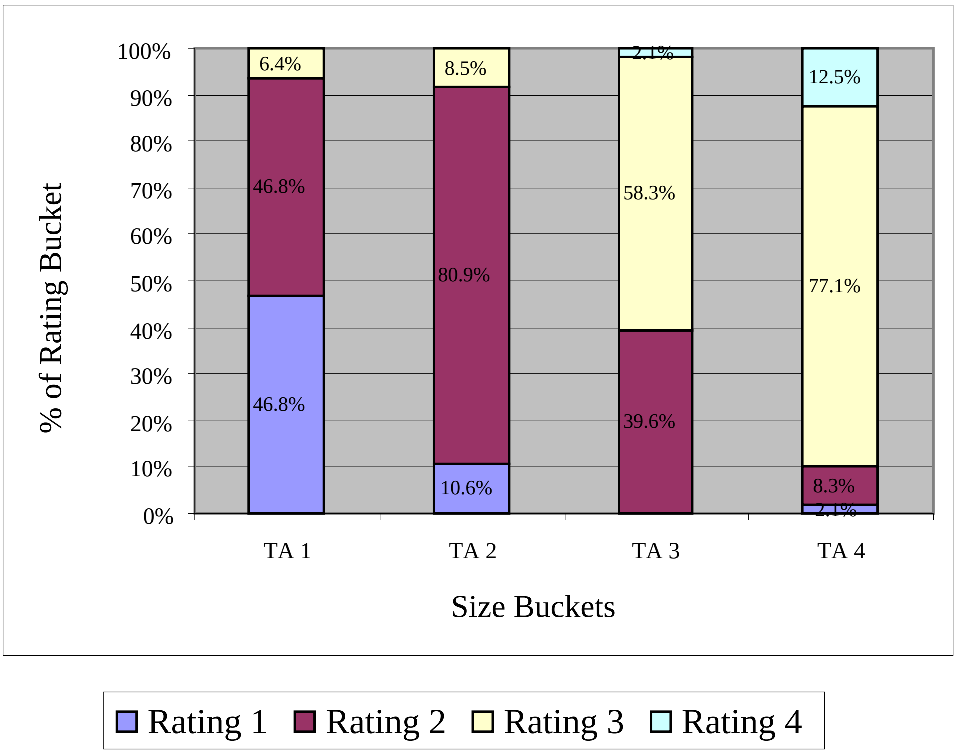

FIGURE 4.7. Relationship between Rating and Size

the rating is extremly dependent on the Total Asset size of the corporate as seen in Figure 4.7 showing size is playing an important role in TA1, TA2, and TA3. But the rating is less dependent on size when the corporate is 'small' between R9 million to R62 million seen in TA4. This will be an important piece of information to remember when we analyse our DEA efficiency results.

4.2. Applying the DEA Framework

4.2.1. Data Segmentation. Having determined the final data set and its characteristics (i.e. the 190 corporates), the data must be prepared and applied to the DEA framework. The first step involves segmenting the data into homogeneous but relevant segments such as industry classification, class, size, rating. Each of these segments will contain sub-segments, for example the segment industry classification contains sub-segments Agriculture, Mining, Construction and so on. We will run corporate data through our DEA model as one large group, but by segmenting the corporate data and running separate sub-segments through our DEA model we will create the ability to delve a little deeper within the initial findings and across various corporate segments.

The first segmentation we will investigate will be by 'class' type with two sub-segments, Listed and Unlisted corporates. We hope to see if this segmentation reveals any further information across the sub-segments. The Unlisted corporates will be further sub-segmented into Unlisted-A and Unlisted-B.

Corporate behavior can be quite diverse across industries. To eliminate this type of diversity we segment the data across industry classifications and run our DEA model on each corporate industry separately. This segmentation will enable us to investigate any findings about corporates within particular industries. In many cases a corporate may operate in more than one industry, where this was the case the most dominant industry is selected by the credit managers i.e. the industry that generates the greatest source of income.

The last segmentation will be across credit ratings. Each sub-segment run through our DEA model will be corporates grouped in the same credit rating. Earlier we saw in Figure 4.4 a large number of corporates clustered around credit ratings 14 and 15. When lending credit, differentiating between corporates is vital, but in this situation where 45% of the corporates are rated either 14 or 15, using the credit rating is futile. We could differentiate using PDs derived directly from a default data model, but as we discussed earlier these PDs were derived by ranking corporates, dividing them into credit ratings and then using default data to map the credit ratings to a median PD. So the granularity of the PD can only go as far as the number of credit ratings. Therefore using the PD is not an option. Another common method used to differentiate within credit ratings, is the use of the raw regression scores produced by the credit rating models to map to a credit rating. This score can only be used if all information, both the quantitative and more importantly qualitative inputs are part of the regression model. If this is not the case, for example the qualitative inputs are overlays to the regression model score, the differentiation is highly likely to become distorted as continuous regression ranges are being modified by discrete values. So if neither the PD nor the regression scores can differentiate between the credit ratings, we should try an alternative. We know that DEA provides efficiency scores produced on a continuous scale from zero to one. If we run each sub-segment of corporates within a credit rating through our DEA model, we could use the DEA efficiency score to rank within credit ratings.

4.2.2. Selecting inputs and outputs. We now determine which corporate information will be used as inputs and outputs into the DEA model. Defining the inputs and outputs is a critical step in the process as this determines what drives the DEA model. The results of the DEA efficiency scores are heavily dependent on this selection.

Within the DEA model, inputs are defined as financial information that is used to drive production of a business. Outputs are defined as financial information that displays the performance of a business. When in doubt we can distinguish inputs from outputs in borderline cases by determining if the driver belongs in the numerator or denominator of the engineering-science efficiency ratio in accordance to whether an increase in this item should result in an increase or a decrease in the efficiency score. For example if we believe a particular driver should increase the efficiency score then it should be represented in the numerator and is most likely an output. If we believe a particular driver should decrease the efficiency score then it should be in the denominator and is most likely an input. Any non-empty subset of inputs and any non-empty subset of outputs may be selected.

Understanding that this step requires experience and knowledge of what drives corporate performance, we consulted a group of credit analysts on the selection of inputs and outputs from the set of financial information available to us. The set of financial information available to us was similar to the financial information used by credit analysts during their day-to-day credit analysis.

As advised by the credit experts, we should always look to include measures from the balance sheet, income statement as well as the cash flow statement. We should also attempt to include information within the following categories: Value/Size, Leverage, Profitability and Growth.

Size and value gives us an idea of what the firm owns and controls with the expectation that it will provide future benefit.

Leverage is related to the extent to which a corporate relies on debt financing, rather than equity. Most corporates use debt to finance their business so this is not necessarily a negative indicator, it becomes negative when there is too much debt which can be measured by thresholds set in financial leverage indicators such as the Debt Ratio (Total Debt/Total Assets), Debt-to-Equity Ratio (Total Debt/Total Equity) and Equity Multiplier (Total asset/Total equity).

Profitability simply measures the extent to which a firm is profitable. The most important conceptual problem with accounting measures of profitability is they do not give a benchmark for making comparisons. A firm is considered profitable economically if it's profitability is greater than investors can achieve on their own in the capital markets. EBITDA (Earnings before interest, taxes, depreciation and Amortization) is a popular metric used to analyze and compare profitability between companies and industries because it eliminates the effects of financing and accounting decisions to evaluate profitability. The only problem with it, is that it is not a standard measure and corporates often change the items included in their EBITDA calculations from one reporting period to another.

The financial information we used was year end statements only. The credit experts stated that ideally we should use the two latest financial years to calculate and incorporate a growth factor (Asset Growth, and Liabilities). Unfortunately this was not possible to obtain for this dissertation.

As a DEA model requires inputs and outputs so too does a credit rating model. When selecting the financial information for inputs and outputs a conscious effort was made to include information that is also used within the ratios that drive the credit ratings within the credit rating models. The difference here is that a credit rating model uses all financial information as inputs, its output is the credit rating.

The DEA model uses financial information as a number of inputs and outputs, the results are efficiencies.

The final set of inputs selected for the DEA model are: Total Assets, as the size of a business is important and the balance sheet item Total Assets is a good indicator of size (Ideally we would have liked to include the current year's and the previous year's Total Assets for a growth function, but this was not available); Total Debt, as corporates generally use debt to finance their business and the balance sheet item Total Debt would capture this; and Total Equity, as this balance sheet item gives us an idea of the reserves and retained profits to date.

The final set of outputs selected of the DEA model are: EBITDA, Total Sales and Funds from Operations (FFO). EBITDA is extracted from the income statement and is used to evaluate profitability. Although Net income is also an indicator of profitability the credit analysts advised us against using this as its results can be skewed. Total Sales also extracted from the income statement indicates the volume of sales. As a rule of thumb,a corporates EBITDA (earnings) should be greater than Total Sales. According to credit experts FFO represents the cash flow from operating activities and is a clear indication of cash on hand.

We are now ready to apply the data to the DEA model and its extensions.

4.2.3. DEA model application. Now that we have analysed our corporate data, segmented it into homogeneous sub-segments, selected a common set of inputs and outputs, we must apply our DEA model. To construct our DEA model several linear programs in equation (4.2.1) must be solved. This must be done in a suitable software package able to compute thousands of variables and constraints simultaneously.

$$
(P_k') \begin{cases} \displaystyle\max_{z^k \in \mathbb{R}^r, w^k \in \mathbb{R}^s} (z^k, Y^k) \\ s.t. \\ \quad (w^k, X^k) = 1 \\ \quad (z^k, Y^{k'}) \leq (w^k, X^{k'}), \qquad k' = 1, 2, \ldots, N. \\ \quad z^k \geq 0, \quad w^k \geq 0. \end{cases} \tag{4.2.1}
$$

The implementation process is simple. Select a sub-segment of corporates. Insert the corporate's inputs, X, and outputs, Y, into the DEA model in equation (4.2.1) and solve. The DEA model will calculate the weights w and z by maximising the weighted sum of outputs over the weighted sum of inputs subject to it's constraints. To determine if a corporate is efficient calculate the cross-efficiency using equation (4.2.2).

$$E_{k,k'} = \frac{(u^k, Y^{k'})}{(v^k, X^{k'})}, \quad k' = 1, 2, \dots, N, \quad k = 1, 2, \dots, N \tag{4.2.2}$$

subject to the constraints

$$u \geq 0, \quad v \geq 0 \tag{4.2.3}$$

$$E_{k,k'} \leq 1 \quad \text{for all } k' \tag{4.2.4}$$

This 'simple efficiency' $E_{k,k}$ can be used to rank the corporates from efficient to least efficient by sorting all the $E_{k,k}$'s in descending order. An alternative ranking can be given by the 'peer efficiency'. The 'peer efficiency', $e_k = \frac{1}{(N-1)} \sum_{k' \neq k} E_{k,k'}$ can be used to rank the corporates by sorting e_k in descending order from most efficient to least efficient.

Where too many corporates appear efficient in the simple efficiency we may *rank order* the efficient corporates using Θ_k^* in equation (4.2.5) in the CWA, SCWA or MCWA methods. To obtain a ranking of most efficient to least efficient among the already efficient corporates we must order Θ_k^* in descending order for any of the three methods CWA, SCWA or MCWA.

$$\Theta_k^* = \frac{\sum_{i=1}^{r} U_i^* y_{i,k}}{\sum_{c=1}^{s} V_c^* x_{c,k}}, \quad k = 1, 2, \dots, N \tag{4.2.5}$$

Having applied the DEA Model, cross-efficiency and ranking we are now ready to analyse the DEA results.

4.3. Results and Analysis

Applying the DEA model to the whole corporate data group we produced simple efficiencies for each corporate. In order to compare simple efficiency scores with credit ratings, the corporates were sorted into quartile efficiencies from the most efficient to the least efficient. Efficiency quartile 1 contains the most efficient corporates and efficiency quartile 4 contains the least efficient corporates. The ratings were also mapped into four rating buckets as seen earlier in Table 4.2 where rating bucket 1 contains all investment grade corporates least likely to default. Rating bucket 2 contains non-investment grade but strong corporates, rating bucket 3 contains average corporates and rating bucket 4 contains weak corporates most likely to default. From this, Figure 4.8 was produced.

Figure 4.8 shows the percentage of corporates within each rating bucket that falls into each efficiency quartile. This analysis uses the entire sample of corporates without distinguishing between them further. If there was no relationship between the efficiency scores and the rating buckets, we would expect roughly 25% of the most efficient corporates in rating bucket 1, 25% of the second efficiency quartile

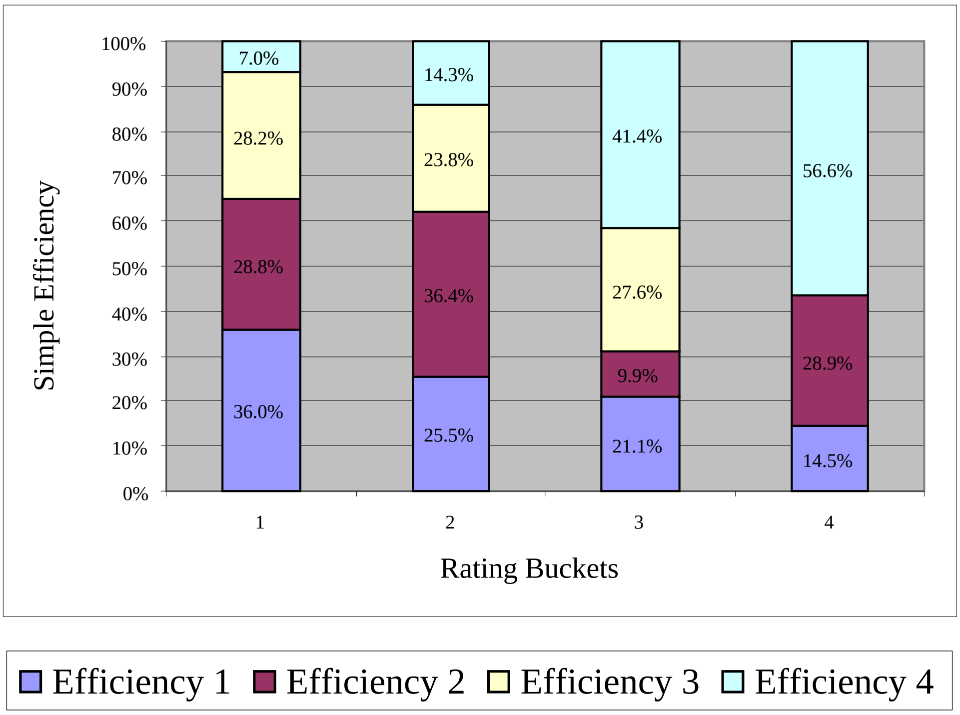

FIGURE 4.8. Comparing Corporate Rating and Simple Efficiency

and so on. However this Figure is displaying an interesting pattern. The most efficient corporates are overrepresented in the strongest rating bucket. In fact 36% of the strongest rated corporates are in the most efficient quartile and 7% are in the least efficient quartile. In contrast we see the opposite in rating bucket 4 with the weakest rated corporates. 14.5% of the weakest corporates are in the most efficient quartile while 56.6% are in the least efficient quartile. There appears to be a trend in Figure 4.8. The percentage of corporates in the most efficient quartile increases as the corporate ratings improve and the percentage of the corporates in the least efficient quartile decreases as the corporate ratings deteriorate. There appears to be no clear pattern in the rating buckets for the 2nd and 3rd efficiency quartiles. Although we can draw no conclusions for the 2nd and 3rd efficiency quartiles, it is clear that the trend we are observing in the 1st and 4th can not be an accident, clearly there is something that both the credit rating and the simple efficiency agree on. The phenomena that we see in Figure 4.8 is similar to what we saw earlier in the U.S. banks literature study in Figure 2.3. Notice the similar pattern when we compare these two Figures.

We also notice there is something that the credit rating and simple efficiency disagree on. Looking at the data from a slightly different perspective, why aren't the strongest rated corporates all efficient and why aren't the weakest rated corporates all inefficient? Notice in Figure 4.8 that if a corporate has a strong rating i.e. rating bucket 1, it is not necessarily always in the most efficient quartile, if a corporate is in a weak rating bucket it is not necessarily always the case that it is in the least efficient quartile. In fact a small percentage of the worst rated corporates appear in the most efficient quartile. Does this make sense? This can only make sense if the efficiency is measuring something different to the credit rating. It is capturing something else in the corporate data that the rating is not measuring. The credit rating by definition is estimating the likelihood of a corporate defaulting. The efficiency is measuring whether a corporate is operating efficiently. We must now ask ourselves if rating and efficiency can be equated because Figure 4.8 is suggesting that although there is some level of consensus, the efficiency has some extra information to offer.

We are implying that a corporate can have the best or the worst rating but this does not mean that they are always operating efficiently. This leads us into many other intriguing questions like: If a corporate is in the better end of the credit rating scale and it is operating inefficiently how does this affect the corporates business in the long term? By operating inefficiently for too long do the corporates start eating into their profits and therefore in time how will it affect their ratings or likelihood to default in the future? Does inefficiency imply that the corporate is likely to migrate to a worse credit rating in the future?

Let us put these questions aside for now and investigate the initial findings further by delving deeper using our segmentation. In our segmentation by class type where we apply our DEA model to separate sub-segments: Listed, Unlisted, with Unlisted further divided into Unlisted-A and Unlisted-B we produce Figures 4.9, 4.10, 4.11.

The data analysis we undertook on the corporate data showed that the Listed corporates made up 24% of the corporate sample data when we counted the number of corporates but when we summed the Total Assets we found it made up 63%. This indicated that the Listed corporates contain the largest corporates. We also observed that the Listed corporates occupy the better end of the credit rating scale compared to the Unlisted-A and Unlisted-B corporates, in fact no Listed corporate is rated worse than credit rating 16 and no Unlisted-B corporate is rated better than 11. In Figure 4.9 and Figure 4.11 it is no surprise that we see an unfilled rating bucket 4 and 1 respectively, based on our data analysis. In Figure 4.9 and 4.11 the percentage of the least efficient quartile increases as the rating deteriorates as we saw in Figure 4.8. In Figure 4.10 the pattern is only there for the first three rating buckets. The last rating bucket 4 shows distorted results. In rating bucket 4 there are too few corporates (two corporates exactly) to create a meaningful bucket size with efficiency distribution. In Figure 4.9 and 4.11 we see the percentage of the 2nd and 3rd efficiency quartile increasing as the rating improves, but other than this no other trends as seen in Figure 4.8 exist. There could be a few explanations for this. The first is that the segmentation by class type is an artificial segmentation, i.e. there is no real difference between the operations of a Listed, Unlisted-A or Unlisted-B corporates. The second has more to do with the number of corporates within each category. Take for example if we divide the corporate group into three categories (Listed, Unlisted-A, Unlisted-B) and then further subdividing these into four categories (Rating bucket 1 to 4) in the best case where all divisions are equal we will get $15 = (190 \div 3) \div 4$ data points in each category. As we see in the Unlisted-A rating bucket 4, these are not equal and this leads to a case where we

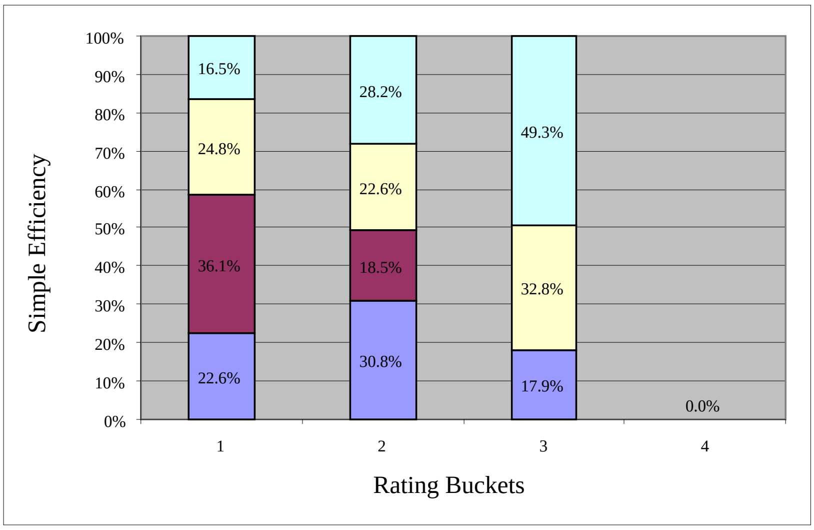

FIGURE 4.9. Comparing Listed Corporate Rating and Simple Efficiency

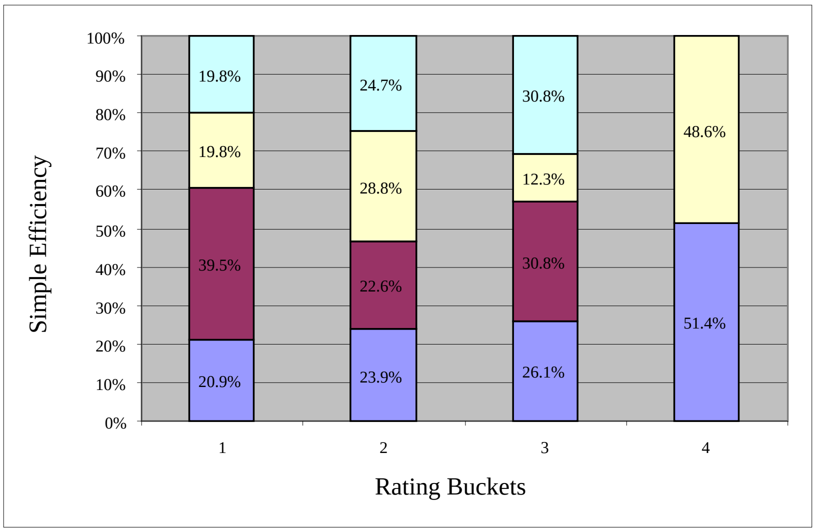

FIGURE 4.10. Comparing Unlisted-A Corporate Rating and Simple Efficiency

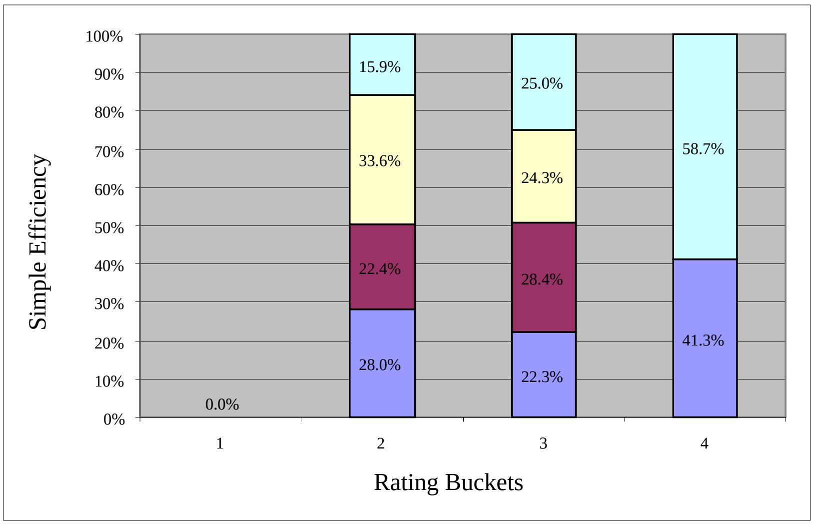

FIGURE 4.11. Comparing Unlisted-B Corporate Rating and Simple Efficiency

Efficiency 1 Efficiency 2 Efficiency 3 Efficiency 4 have only two corporates used to assess the relationship between efficiency and rating in one rating bucket which is too small to use as a sample.

Knowing that segmentation may present some problems it is still worthwhile turning our attention to corporates within industries. Corporate behavior can be quite diverse across industries, segmentation within industry classification will enable us to investigate and obtain insight into corporates within a particular industry. In Figure 4.12, 4.13, 4.14, 4.15 and 4.16 we compare simple efficiency and peer efficiency to the credit ratings using least squares regression analysis. The objective is to find any meaningful trend or relationship between the credit rating and efficiency within industries.

Let us investigate some of the industries we highlighted earlier. Retail General had the largest number of corporates but was relatively small when we summed their Total Assets. The regression results show no clear trend, both the simple efficiency and the peer efficiency show a trend line which is quite close to flat. In the data analysis we investigated in Table 4.1 we also notice that most of these corporates are not spread evenly but clustered in the centre of credit rating scale. Telecommunications, this industry occupies the better end of the rating scale as seen in Figure 4.1. This industry has been 'booming' over the last few years and the credit managers have clearly identified this and placed three of the five corporates in investment grade credit ratings. Along with the corporates operating in the Oil industry, these corporates do not have a rating worse than 14, so we can safely say the Telecommunication and Oil appear to be the top performing industries. The DEA model's simple efficiency results in Figure 4.12, 4.13, 4.14, 4.15 and 4.16 interestingly identifies these industries as being top performers and all the corporates within them achieve the best efficiency of one. The Mining industry, a dominant industry when we look at sum of Total Assets, has a positive trend i.e. as the credit rating deteriorates the efficiency improves, contrary to what we would expect. The Chemicals industry with the second highest sum of Total Assets has a negative trend as we would expect. It is also encouraging to see that in this industry there is one corporate that according to the credit rating outperforms the group by 5 grades but is also identified as efficient in the simple efficiency results. The peer efficiency in these figures does not appear to be giving us much more information, in fact it appears that the peer efficiency is creating an averaging effect when each of the corporates get to rate each other. Having seen the results in Figure 4.8 we would naturally expect negative trends in the figures i.e. where the rating improves the efficiency also improves. In the industry regressions there is roughly an equal number of industries with a positive and negative trend, so no conclusions can be drawn between the credit rating and the efficiency within industries.

Earlier we discussed how DEA could be used to rank corporates within credit ratings where no further differentiation could be made using the PD or the raw credit rating model regression score. In Table 4.3 and 4.4 below we have used DEA results to rank the corporates within credit ratings 14 and 15 as they contain over 45% of the portfolio and would clearly pose the biggest problem if further differentiation was required. Both the simple efficiency or the peer efficiency can be used to rank order, there is very little difference between the two. The peer efficiency can be used if it is more appropriate to view each corporate from the perspective of

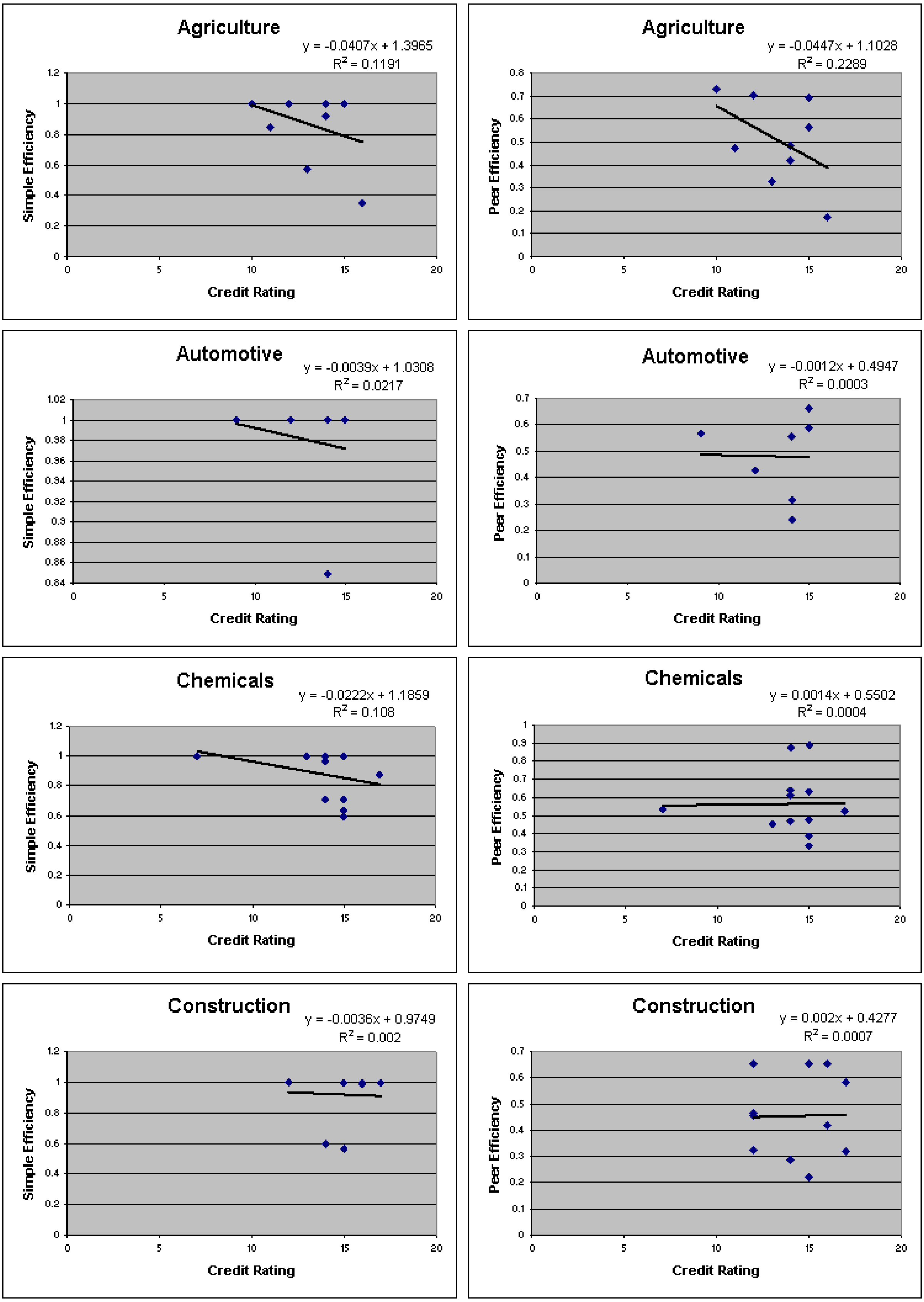

FIGURE 4.12. Application of DEA to Corporate Industries

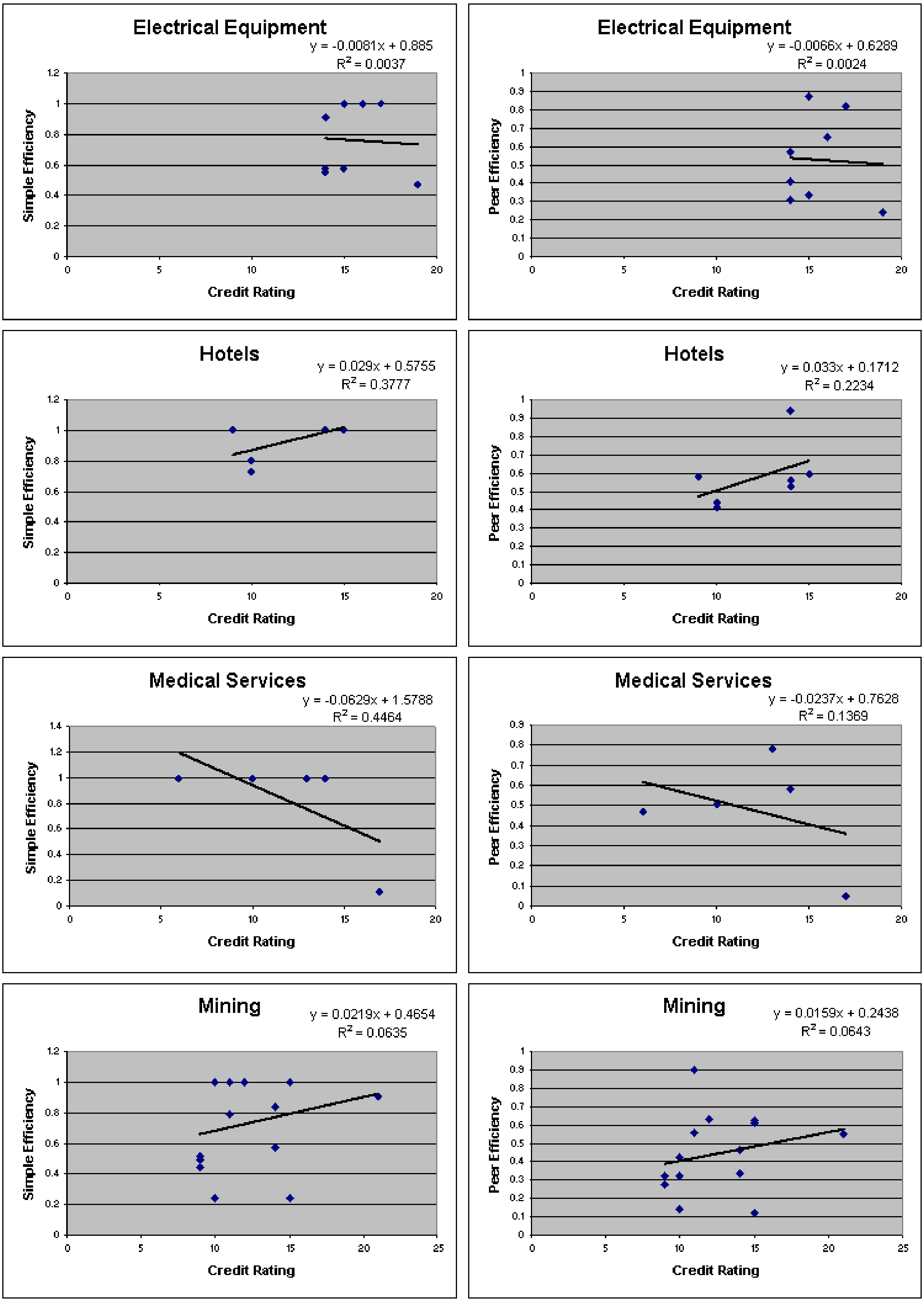

FIGURE 4.13. Application of DEA to Corporate Industries

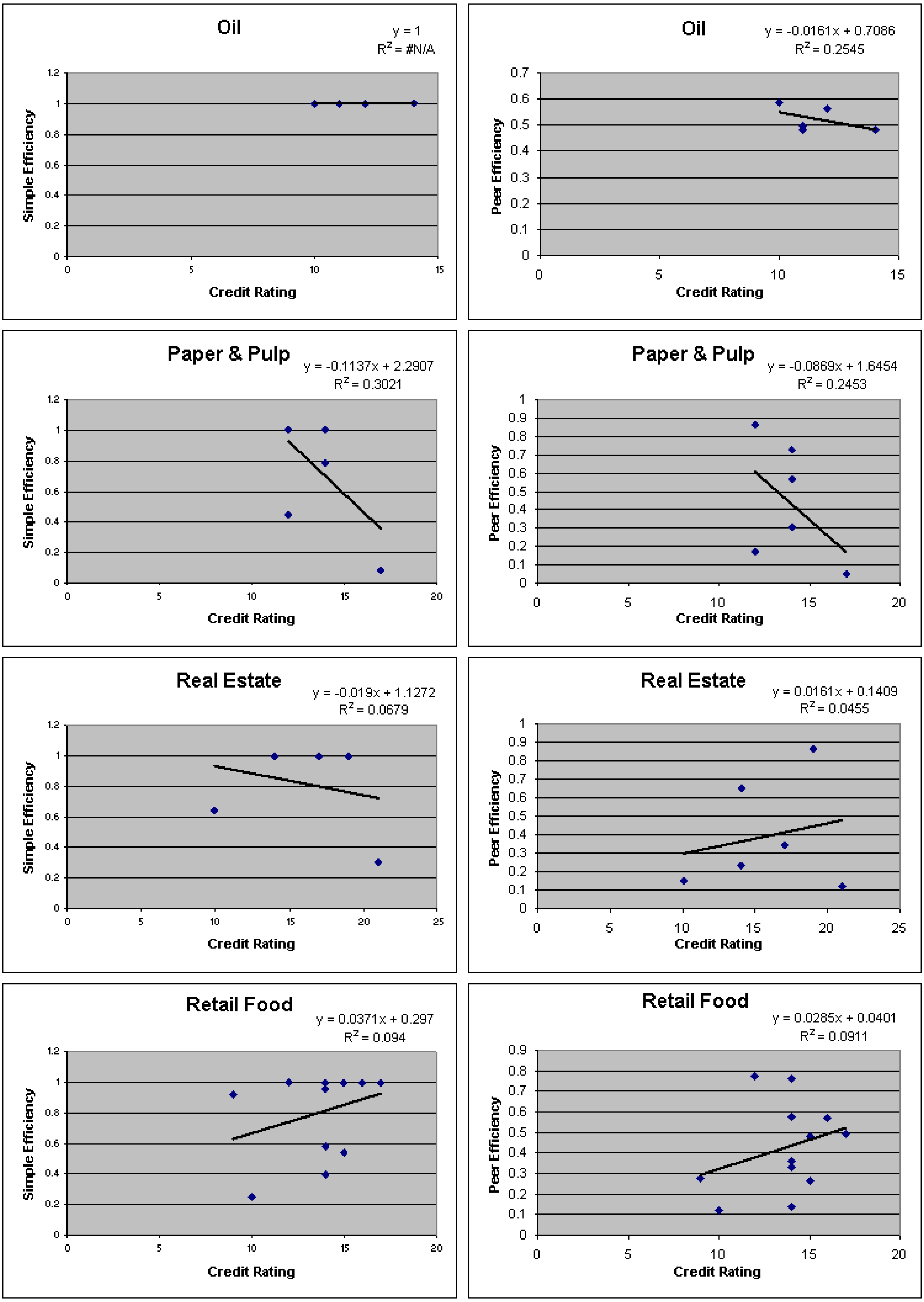

FIGURE 4.14. Application of DEA to Corporate Industries

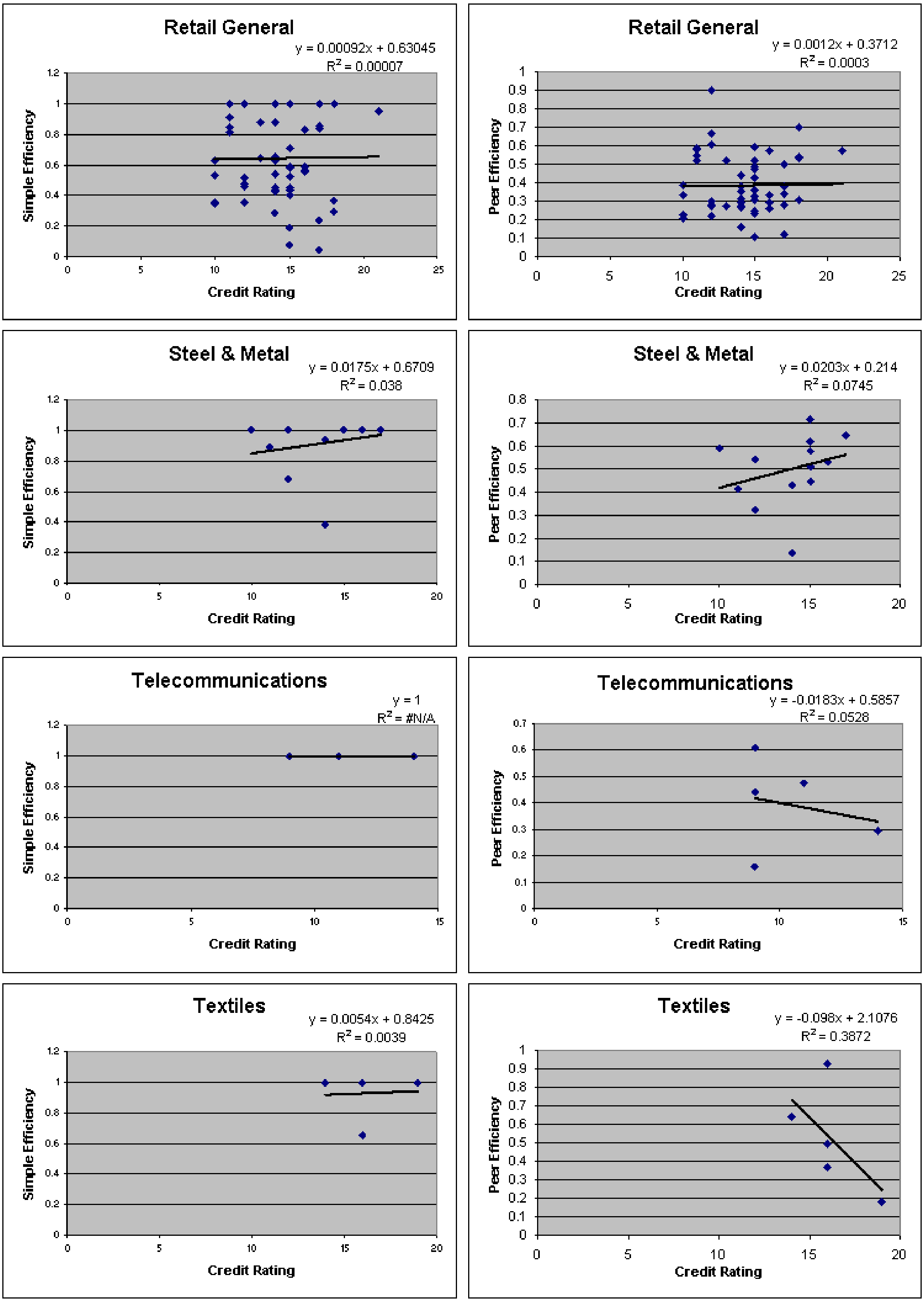

FIGURE 4.15. Application of DEA to Corporate Industries

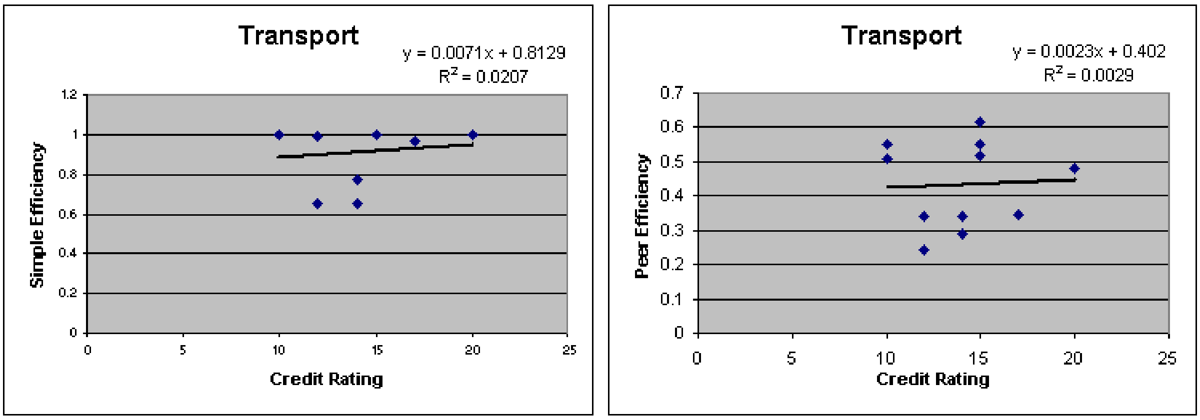

FIGURE 4.16. Application of DEA to Corporate Industries

the rest of the group or if too many corporates appear efficient and differentiation is needed for the efficient DMUs. We have included the maverick indicators as they generally highlight where the simple efficiency and the peer efficiency differ.

If it is more appropriate to use the simple efficiency and we would like to differentiate between the efficient DMUs we can use methods CWA, SCWA or MCWA [**12**]. We will now apply the three models CWA, SCWA and MCWA methods and calculate the peer efficiency for two sub-segments, rating 16 where ten corporates were classified as efficient and the Steel & Metal industry where nine corporates were classified as efficient. Note we are only using these methods for efficient DMUs. In Table 4.5 we find that although the rankings vary slightly for CWA, SCWA and MCWA they agree on the first two (C071 and C058) and the last corporate C263. The peer efficiency which is a ranking made by peers using cross-efficiencies shows major differences to the other three ranking models. Corporate C263 is showing very interesting results. According to C263's peers in rating 16 (including the inefficient corporates), C263 is regarded highly. C263 is receiving a very high peer appraisal and it is not a maverick. In Table 4.6, the Steel & Metal industry, we see the same rankings for the last two corporates (C274 and C338) and similar agreement in the top end of the rankings. Although we do see all four methods have ranked C192 in the top three spots the peer efficiency is showing very different results.

Let us investigate each input and output used in the DEA model and compare them to the Efficiency and the Credit Rating. In Figure 4.17 and 4.18 we have created buckets which contain an equally sized number of corporates for inputs, Total Assets, Total Debt, Total Equity, and outputs EBITDA, Total Sales and Funds from Operations, where bucket 1 contains the corporates with the largest input or output and bucket 4 contains the corporates with the smallest input or output. On the left hand side we have the percentage of efficiency quartiles in each bucket and on the right hand side we have the percentage of rating buckets. Focusing our attention on Total Assets notice the differences and similarities of the left and

ID	Industry	Credit Rating	Class	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)
C343	Agriculture	14	Listed	1	0.503	0.987
C035	Automotive Parts	14	Unlisted-A	1	0.489	1.044
C339	Electricity Equip	14	Unlisted-B	1	0.560	0.787
C081	Hotels	14	Listed	1	0.465	1.152
C294	Hotels	14	Unlisted-A	1	0.796	0.257
C014	Medical Services	14	Unlisted-A	1	0.607	0.648
C176	Oil	14	Unlisted-B	1	0.423	1.363
C296	Real Estate	14	Listed	1	0.252	2.973
C152	Pulp & Paper	14	Unlisted-B	1	0.740	0.352
C264	Steel & Metals	14	Unlisted-B	1	0.664	0.506
C034	Retail General	14	Unlisted-B	1	0.718	0.393
C100	Retail General	14	Unlisted-B	1	0.775	0.291
C234	Retail General	14	Unlisted-B	1	0.528	0.894
C254	Retail General	14	Unlisted-B	1	0.552	0.811
C159	Retail General	14	Unlisted-A	1	0.651	0.536
C304	Retail Food	14	Unlisted-A	1	0.797	0.255
C337	Chemicals	14	Listed	0.885	0.630	0.405
C013	Hotels	14	Unlisted-A	0.880	0.443	0.984
C305	Retail Food	14	Unlisted-A	0.866	0.611	0.416
C266	Retail General	14	Unlisted-B	0.837	0.435	0.925
C054	Transportation	14	Unlisted-B	0.802	0.508	0.579
C171	Retail General	14	Listed	0.782	0.523	0.496
C039	Retail General	14	Unlisted-A	0.754	0.480	0.571
C041	Chemicals	14	Unlisted-A	0.727	0.448	0.624
C132	Chemicals	14	Unlisted-A	0.703	0.418	0.679
C008	Agriculture	14	Listed	0.684	0.438	0.563
C200	Pulp & Paper	14	Unlisted-A	0.675	0.348	0.939
C307	Electricity Equip	14	Unlisted-A	0.669	0.457	0.463
C184	Pulp & Paper	14	Unlisted-A	0.638	0.350	0.825
C045	Automotive Parts	14	Unlisted-A	0.635	0.453	0.403
C046	Textiles	14	Unlisted-A	0.612	0.428	0.430
C018	Transportation	14	Unlisted-A	0.591	0.293	1.017
C093	Chemicals	14	Unlisted-A	0.545	0.315	0.731
C060	Retail Food	14	Unlisted-A	0.531	0.274	0.934
C066	Retail Food	14	Unlisted-A	0.517	0.345	0.498
C016	Construction	14	Unlisted-B	0.501	0.295	0.697
C299	Electricity Equip	14	Unlisted-A	0.491	0.300	0.634
C300	Real Estate	14	Unlisted-B	0.489	0.300	0.628
C218	Steel & Metals	14	Unlisted-A	0.486	0.210	1.317
C331	Mining	14	Unlisted-A	0.475	0.291	0.633
C026	Retail Food	14	Unlisted-A	0.447	0.194	1.303
C166	Retail General	14	Unlisted-A	0.423	0.279	0.518
C121	Automotive Parts	14	Unlisted-A	0.409	0.191	1.143
C193	Mining	14	Listed	0.404	0.213	0.897
C157	Telecommunications	14	Listed	0.322	0.158	1.037

TABLE 4.3. Differentiating within Credit Rating 14 using DEA

ID	Industry	Credit Rating	Class	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)
C215	Electricity Equip	15	Unlisted-B	1	0.506	0.978
C155	Steel & Metals	15	Unlisted-B	1	0.501	0.996
C002	Transportation	15	Unlisted-B	1	0.585	0.710
C135	Transportation	15	Unlisted-B	1	0.753	0.327
C293	Transportation	15	Unlisted-A	1	0.805	0.243
C120	Retail General	15	Unlisted-B	1	0.772	0.296
C074	Retail Food	15	Unlisted-A	1	0.506	0.977
C174	Retail General	15	Unlisted-B	0.999	0.442	1.263
C028	Retail General	15	Unlisted-B	0.918	0.532	0.725
C143	Retail General	15	Unlisted-B	0.882	0.439	1.009
C324	Construction	15	Unlisted-B	0.789	0.481	0.642
C186	Retail General	15	Listed	0.728	0.440	0.652
C083	Retail General	15	Unlisted-B	0.728	0.372	0.956
C273	Mining	15	Listed	0.686	0.400	0.716
C338	Steel & Metals	15	Unlisted-B	0.664	0.426	0.558
C010	Steel & Metals	15	Unlisted-A	0.652	0.335	0.946
C087	Mining	15	Unlisted-A	0.630	0.354	0.777
C230	Retail General	15	Listed	0.604	0.313	0.931
C110	Retail General	15	Unlisted-B	0.570	0.387	0.474
C172	Retail General	15	Unlisted-A	0.555	0.262	1.117
C244	Chemicals	15	Unlisted-B	0.554	0.374	0.483
C252	Steel & Metals	15	Unlisted-B	0.543	0.342	0.587
C095	Automotive Parts	15	Unlisted-B	0.528	0.301	0.752
C105	Retail General	15	Unlisted-A	0.482	0.318	0.517
C270	Retail General	15	Unlisted-B	0.451	0.216	1.085
C286	Agriculture	15	Unlisted-A	0.440	0.287	0.532
C044	Agriculture	15	Unlisted-B	0.418	0.278	0.501
C258	Retail General	15	Unlisted-B	0.408	0.242	0.683
C040	Chemicals	15	Unlisted-A	0.382	0.187	1.048
C335	Hotels	15	Unlisted-B	0.371	0.239	0.553
C144	Retail Food	15	Unlisted-B	0.346	0.217	0.590
C303	Chemicals	15	Unlisted-B	0.323	0.163	0.981
C131	Electricity Equip	15	Listed	0.285	0.189	0.505
C232	Chemicals	15	Unlisted-B	0.263	0.127	1.079
C287	Chemicals	15	Unlisted-B	0.243	0.123	0.978
C260	Construction	15	Listed	0.205	0.135	0.521
C037	Automotive Parts	15	Unlisted-B	0.187	0.273	-0.315
C192	Steel & Metals	15	Unlisted-B	0.166	0.506	-0.673
C079	Retail General	15	Unlisted-B	0.165	0.099	0.655
C241	Mining	15	Unlisted-A	0.161	0.093	0.739

Table 4.4. Differentiating within Credit Rating 15 using DEA

Rank	CWA			SCWA			MCWA				Peer efficiency		
	DMU	Θ_k^{C*}	Δ_k^{+*}	DMU	Θ_k^{S*}	Δ_k^{-*}	DMU	Θ_k^{M*}	Δ_k^{+*}	Δ_k^{-*}	DMU	e_k	Maverick
1	C071	1	0	C071	2.97	0.74	C071	1.67	0	0.33	C263	0.81	0.23
2	C056	0.79	0.08	C056	2.80	0.27	C056	1.25	0	0.06	C007	0.68	0.47
3	C240	0.69	0.18	C084	1.98	0.14	C240	1.13	0	0.04	C187	0.57	0.75
4	C187	0.63	0.66	C240	1.88	0.25	C187	1	0	0	C056	0.52	0.93
5	C007	0.56	0.11	C007	1.57	0.07	C084	0.93	0.01	0	C245	0.51	0.95
6	C084	0.56	0.16	C187	1.55	0.56	C007	0.90	0.02	0	C178	0.48	1.08
7	C211	0.39	0.14	C178	1.02	0	C245	0.62	0.06	0	C240	0.47	1.13
8	C245	0.39	0.15	C211	1.02	0	C211	0.62	0.05	0	C071	0.47	1.14
9	C178	0.34	0.19	C245	1	0	C178	0.55	0.08	0	C084	0.42	1.38
10	C263	0.26	0.18	C263	1	0	C263	0.43	0.09	0	C211	0.38	1.64
	$\Delta^* = 1.859$			$\Delta^* = 2.025$			$\Delta^* = 0.75$						
Weights	$U_1 = 1, U_2 = 1, U_3 = 1$ $V_1 = 4.916, V_2 = 1, V_3 = 1$			$U_1 = 1, U_2 = 1.423, U_3 = 1$ $V_1 = 1, V_2 = 1, V_3 = 3.317$			$U_1 = 1, U_2 = 1, U_3 = 1$ $V_1 = 2.774, V_2 = 1, V_3 = 1$						

TABLE 4.5. Comparison between CWA, SCWA, MCWA and Peer Efficiency for Rating 16's efficient DMUs

Rank	CWA			SCWA			MCWA				Peer efficiency		
	DMU	Θ_k^{C*}	Δ_k^{+*}	DMU	Θ_k^{S*}	Δ_k^{-*}	DMU	Θ_k^{M*}	Δ_k^{+*}	Δ_k^{-*}	DMU	e_k	Maverick
1	C155	1	0	C198	4.12	1.04	C198	1.88	0	0.26	C192	0.71	0.40
2	C198	1	0	C192	2.59	0.45	C192	1.31	0	0.07	C274	0.64	0.55
3	C192	0.97	0.02	C253	1.82	0.19	C155	1	0	0	C155	0.62	0.62
4	C010	0.91	0.27	C252	1.65	0.71	C252	1	0	0	C198	0.59	0.70
5	C252	0.90	0.18	C155	1.64	0.51	C010	1	0	0	C338	0.58	0.73
6	C071	0.79	0.13	C010	1.41	0.82	C253	0.99	0	0	C253	0.54	0.85
7	C253	0.66	0.15	C071	1.12	0.04	C071	0.69	0.09	0	C071	0.53	0.89
8	C274	0.38	0.24	C274	1	0	C274	0.64	0.06	0	C252	0.50	0.98
9	C338	0.31	1.15	C338	1	0	C338	0.45	0.41	0	C010	0.44	1.26
	$\Delta^* = 2.136$			$\Delta^* = 3.742$			$\Delta^* = 0.893$						
Weights	$U_1 = 11.383, U_2 = 1, U_3 = 1$ $V_1 = 3.074, V_2 = 1, V_3 = 1$			$U_1 = 1, U_2 = 2.770, U_3 = 1$ $V_1 = 1, V_2 = 6.271, V_3 = 1$			$U_1 = 1.883, U_2 = 1, U_3 = 1.225$ $V_1 = 1, V_2 = 1.444, V_3 = 1$						

TABLE 4.6. Comparison between CWA, SCWA, MCWA and Peer Efficiency for Industry Steel & Metal's efficient DMUs

right hand side. There are a couple of items which should be highlighted. Starting with the similarities between efficiency and rating, we see that as Total asset size decreases the percentage of the most efficient corporates decreases, the least efficient corporates increase. A similar pattern takes place for rating. Generally as Total Assets decreases the likelihood of a corporate being less efficient and having a worse rating increases. However there is a very big difference in how this is taking place when we compare efficiency and rating. Notice how in each of the input and output buckets there is no one efficiency quartile that dominates any of the input and output buckets.

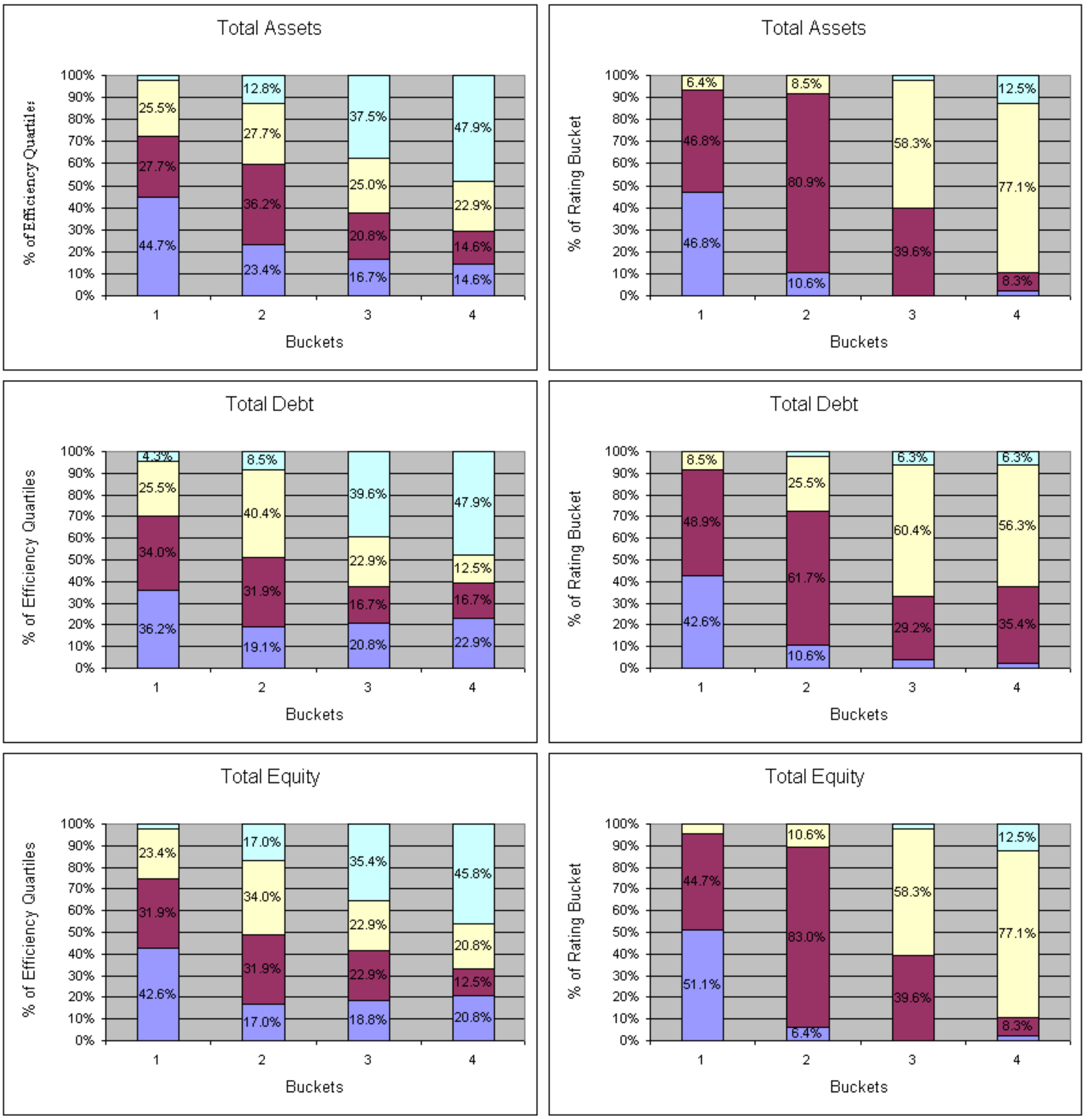

FIGURE 4.17. Corporate Inputs used in DEA model and Credit Rating

Turning our attention to the Rating we see this is not the case, Total Assets bucket 2 is dominated by rating bucket 2, Total Assets bucket 4 is dominated by rating bucket 3. This appears to indicate that size is dictating very clearly which rating bucket a corporate belongs. Looking at the other inputs and outputs, this dominance in % of rating bucket in input and output bucket's 2 and 4 is appearing in every graph. Also notice how little of rating bucket 1 and rating bucket 4 is represented in any of the rating bucket graphs. It is clear that there is a lack of

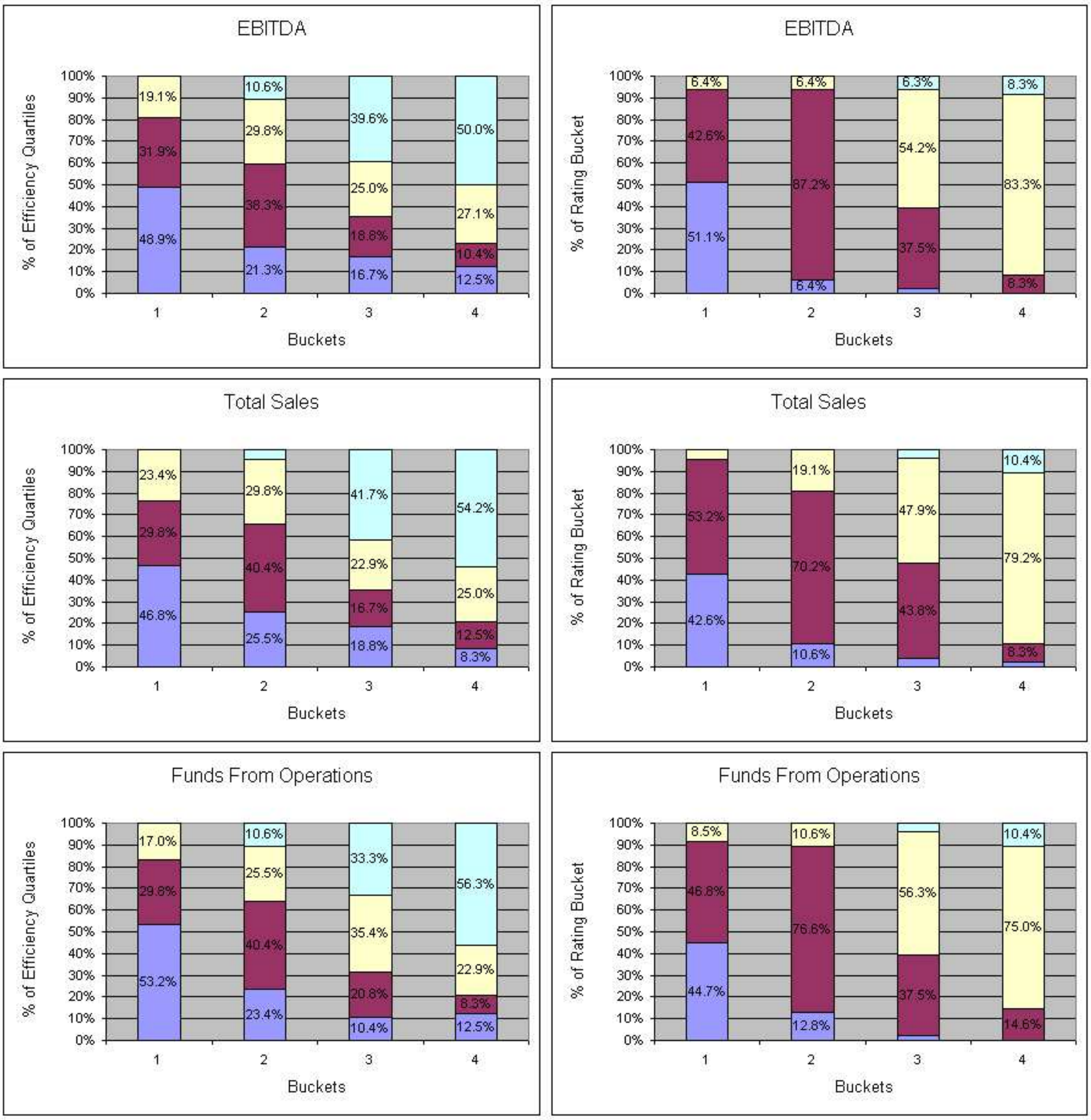

FIGURE 4.18. Corporate Outputs used in DEA model and Credit Rating

differentiation here and rating bucket 2 and 3 take over input and output buckets 2 to 4 in every case. Remember in our data analysis when we discussed that over 45% of the corporate group is given a credit rating of 14 or 15, the same picture is playing out in these graphs. We can now understand why the ratings are not supplying the differentiation needed between corporates as size dominates the scene.

Turning to the efficiency notice that in nearly all the input an output buckets each efficiency quartile is represented and although there is a general trend that as the buckets size decreases from 1 to 4 and the least efficient quartiles increases, they do not dominate the buckets providing more differentiation between the corporates in each input and output and therefore overall.

4.4. Conclusion

In this chapter we applied DEA to corporate entities and obtained DEA results in the form of simple efficiencies, peer efficiencies, mavericks and corporate rankings. We investigated the data characteristics. One potential problem we identified with the data upfront, was that the internal ratings were heavily clustered around the centre of the rating scale. Over 45% of the corporates are rated 14 or 15. According to the Basel II requirements this is unfavourable and further differentiation is necessary. When comparing the credit rating to the size of the corporate we found that if the corporate is 'large' between R62 million to R88 billion the credit rating is extremely dependent on the size of the corporate, but the rating is less dependent on size when the corporate is 'small' between R9 million to R62 million.

We applied our DEA model to the whole corporate group and corporate sub-segments across class type, industry classification and credit rating. For the whole corporate group, the results showed similar trends between simple efficiency and credit rating. The percentage of corporates in the most efficient quartile increased as the corporate ratings improved and the percentage of the corporates in the least efficient quartile decreased as the corporate ratings deteriorated. This trend confirms the commonsense notion that corporates that receive better credit ratings from the credit analysts are more efficient. This result points to the potential usefulness of the DEA model.

Although there is a trend between efficiency and credit rating this does not mean that the corporates in the best credit rating are always in the most efficient quartile. In fact some of the corporates in the best credit rating were found to be inefficient. This leads us to believe that efficiency gives extra information about these corporates over and above the credit rating. In this case, the corporate may not be risky but the DEA model recognises that they are not utilising their resources to their full potential and are operating inefficiently. If we had the opportunity to investigate the migration of ratings of the efficient and inefficient corporates through time this would be useful. It can help determine if inefficiency can affect a corporate's business and therefore credit rating through time. Unfortunately, migration on a clustered credit rating scale where 45% of the corporates lie in two grades would show very little meaningful migration. This exercise would need to be performed on a data set which is far more evenly distributed across the credit rating scale.

We tried to delve a little deeper by running the DEA model on separate class type sub-segments (Listed, Unlisted-A, Unlisted-B corporates) and by separate industry sub-segments but we did not gain any further insight into the corporate efficiencies and credit ratings relationship. This was due to the artificial segments and the common segmentation trap where too few or no data points filled categories within each sub-segment.

We compared the inputs and outputs to the DEA model results and the credit rating. We found that the efficiency was more or less evenly distributed with a decreasing trend as the input or output decreased, but this was not the case for the rating. Although the trend was present there were rating buckets that dominated most of the inputs and outputs buckets, re-emphasizing the clustering problem and lack of differentiation seen in the credit rating that we picked-up in our raw data.

With the problematic clustering of the credit ratings in 14 and 15 we decided based on the Basel II requirements that differentiation between corporates is necessary for corporate lenders. Where the PD and or the credit rating model score can not be used to differentiate between corporates we demonstrated how DEA can provide this granularity very easily using either simple efficiency or peer efficiency. This demonstrates that DEA is a tool that facilitates banks meet the Basel II requirements. We also demonstrated the ranking of efficient corporates in more recent DEA development methods of CWA, SCWA and MCWA methods for the Steel & Metal industry and credit rating 16 where many corporates appeared efficient.

In summary simple efficiency and credit rating have some elements in common and DEA can play an important role in credit assessment as a supplementary tool to credit rating models to help meet the Basel II requirements. We confirmed that corporates receiving better credit ratings are generally more efficient, but the efficiency is providing more information about the corporate. Credit rating and efficiency are identifying and measuring two different items. Credit rating is measuring the risk of the client and its likelihood of default. Efficiency is purely interested in the operations of the corporate under its current conditions and whether it is utilising it resources effectively to produce more output compared to every other corporate. Where results appeared inconclusive, this was attributed to the clustering of the credit ratings creating serious data challenges. However this clustering did help us demonstrate how easy it is to apply DEA and produce differentiation between corporates using efficiency in single credit ratings. Further research entails applying DEA to a corporate data sample which is more evenly distributed across credit ratings and investigating relationships between efficiency and corporate credit rating migration.

5

Application to Countries

In this chapter we will apply DEA to country risk and obtain country efficiency scores as well as country rankings. We are again interested in differences or relationships between the DEA country efficiency results and country credit ratings. Where country ratings do not provide enough granularity we will apply DEA to rank the countries based on efficiency.

5.1. Sample Data

5.1.1. Data Collecting and Cleaning. Ideally we would have liked to collect data from as many countries as possible from around the world. Unfortunately finding free, reliable and comparative economic statistics across different countries proved to be a mammoth task. For some developing countries the economic statistics that are available are not reliable. For the developed countries many have different policies and slight calculation variations when producing commonly known economic statistics, making comparison between countries impossible. In order to reduce the problems of data reliability and comparability, we sourced our country data from an organisation that has made it its task to provide comparative and reliable economic statistics in the countries where this information can be collected.

This organisation is known as the OECD. The country data utilised was sourced from OECD publicly available published data. OECD stands for Organisation for Economic Co-operation and Development. It is a Paris-based club for industrialised countries ('the best'). It was formed in 1961, building on the OEEC (Organisation for European Economic Co-operation). By 2003 the OECD groups 30 member countries sharing a commitment to democratic government and the market economy. Together the OECD countries produce two-thirds of the world's goods and services [**13**].

OECD is best know for its publications and statistics, its education, development and science and innovation. We have sourced country comparisons of main economic indicators from 30 OECD countries for years 2004 to 2006, from the OECD in April 2007. The data can be found in [**15**]. The OECD produce country risk classifications also known as country risk ratings. These measure a country's credit risk. i.e. the likelihood that a country will service its external debt. There are eight (0 to 7, where 0 is the best credit quality) country risk categories, derived through a model which uses quantitative and qualitative assessments. The final rating assigned to the countries is a consensus of a group of country risk experts which meet several times a year. The final list of country risk ratings are then published after each meeting [**13**]. OECD country ratings were collected for 2004 to 2007.

The country data contains one of the most comprehensive and comparative lists of macroeconomic information for the 30 OECD countries. The macroeconomic indicators published included: Gross Domestic Product (GDP), Private Consumption, Government Consumption, Gross Fixed Capital Formation (GFCF), Exports and Imports from the National Account, Industrial Production, Passenger Car Registration, Retail Trade, Hourly Earnings, Unit Labour Costs, Standardised Unemployment Rates, Civilian Employment, Short-Term and Long-Term Interest Rates, Share Prices, Monetary Aggregates, Real Effective Exchange Rates, Exchange Rates and International Trade of Imports and Exports. As will be seen in the sections to follow not all the economic indicators will be relevant for our DEA model and this will be discussed in more detail later along with a detailed data analysis of the economic indicators that are selected. Like every data source the data contained a few missing values. Where a large portion of country or economic indicator data was missing they were excluded from the final data set. After data cleaning, the final data set contained 25 OECD countries with their macroeconomic indicators for 2004 to 2006 and corresponding OECD country ratings for 2004 to 2007.

5.2. Applying the DEA Framework

5.2.1. Selecting inputs and outputs. By interrogating the macroeconomic indicators available to us after data cleaning, we can determine which main economic indicators should be classified as inputs and outputs into the DEA model. Defining the inputs and outputs is a critical step in the process as this determines what drives the DEA model. The results of the DEA efficiency scores are heavily dependent on this selection.

Within our DEA model inputs are defined as economic indicators that drive the production of a country. Outputs are defined as economic indicators that display the performance of a country. To be able to identify which main economic indicators are inputs and outputs into our DEA model, let us get a better understanding of the indicators sourced from the OECD.

Gross Domestic Product (GDP) is one of the ways for measuring the size of a countries economy. It is the market value of all the goods and services producted by labour and property located in a country. Simply, GDP = consumption + investment + (government spending) + (exports - imports). From this description we can clearly see that GDP can only be an output as it is displaying the performance of a country.

Private Consumption includes most personal expenditures of households such as food, rent, medical expenses and so on, but does not include new housing. Government Consumption covers spending of goods and services (defense, judicial system, education, etc); it therefore excludes the components of government spending that are transfers (such as social security and unemployment benefits). It is usually counter-cyclical as it tends to increase during recessions. Government spending tends to be a constant share of GDP and is not very volatile. More than spending, it is the budget deficit, the difference between total government spending and revenues, that may affect asset prices. Consumption (both public and private) can be considered an input as it drives production.

Gross Fixed Capital Formation (GFCF) functions as an economic indicator of the level of business activity. It is usually defined as the total value of additions to fixed assets by resident producer enterprises, less disposals of fixed assets during the quarter or year, plus additions to the value of non-produced assets (such as discoveries of mineral deposits, or land improvements). GFCF can only be an output, as the definition clearly states it indicates a country's level of business activity.

Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude labour and property income (formerly called factor services) as well as transfer payments. Where exports of goods are valued f.o.b., the costs of transportation and insurance up to the border of the exporting country are included in exports of goods. We notice that there are two economic indicators for exports in sourced OECD data. The first is named 'Exports of Services and Goods - National Account' and the second is 'International Trade (Exports)' The first indicator is based on each country's own system of national accounts. The second comes from national statistical institutes, e.g Eurostat for the Euro area countries. From this description exports show the performance of the country, therefore it will be considered as an output. For our DEA model we don't need to double account for export and therefore selecting a single export indicator will suffice. There appears to be little difference between the two export indicators therefore for no particular reason we will select the International Trade (Exports) indicator as an output into our DEA model.

Imports of goods and services represent the value of all goods and other market services received from the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude labour and property income (formerly called factor services) as well as transfer payments. Similar to exports there are two indicators for imports. The first 'Imports of Goods and Services - National Account' based on each country's own system of national accounts, the second comes from national statistical institutes, e.g Eurostat for the Euro area countries. Again we do not need two import indicators so we will select 'International Trade (Imports)' and it will be considered an input as imports appear to drive a country instead of display the country's performance.

Industrial production is the total produce of factories and mines within a country. The OECD use an index of industrial production (IIP) as the reference series for aggregate economic activity because it constitutes the most cyclical subset of the aggregate economy and is available promptly for most OECD countries [**14**]. Industrial production is clearly an output, but the OECD have stated in [**14**] that the cyclical profiles of the IIP and GDP in OECD countries have been found to be closely related. With Industrial Production having such a close relationship to GDP, we should only include one of these economic indicators in our DEA model. Since we have included Private Consumption, Government Consumption, Imports and Exports which are all used in calculating GDP let us exclude GDP as it is already present in it's make-up and rather include Industrial Production as an output into our DEA model.

Hourly Earnings are the seasonally adjusted hourly earnings within the manufacturing industry paid per employed person per hour, including overtime pay and regularly recurring cash supplements. Unfortunately [**14**] warns that this definition varies from country to country, particularly with respect to workers covered, treatment of bonuses, retrospective wage payments and size of reporting unit. [**14**] lists so many differences that for our purposes we will exclude this indicator as it may introduce noise, rather than contribute to our final results.

Standardised Unemployment rates are the numbers of unemployed persons as a percentage of the civilian labour force. Civilian labour force is defined as the sum of the unemployed plus those in civilian employment. Civilian employment is derived from labour force sample surveys collected and seasonally adjusted by the OECD [**14**]. Let us discuss the possibility of including one of these indicators (Employment or Unemployment) as they are exact opposites of each other. Employment is a tricky macroeconomic indicator. How we decide if this is an input or an output really depends on whether the country's objective is to create jobs or not. Let us assume that the objective is to create jobs then, employment would be a target (output) which is either met or not met. Therefore in this situation treating Civilian Employment as an output makes sense. Not all countries treat employment as a target and comparison of this indicator across countries becomes ambiguous, so for this reason we will exclude this indicator in our DEA model.

Retail Trade initially appeared to be an excellent indicator to include in our DEA model because it measures the volume of retail trade within each country. It was found that in some cases the volumes used for different countries were calculated on different goods and services and [**14**] warned that care should be taken when using this indicator particularly across countries. For this reason we will exclude this indicator in our DEA model.

Interest Rates, Exchange Rates, Unit Labour costs and other economic indicators were either excluded because of their relevance, comparability problems or lack of data supplied for the 25 OECD countries.

After this detailed interrogation of all economic indicators sourced from the OECD, the final set of inputs and outputs that will be used in our DEA model can be seen in Table 5.1.

MAIN ECONOMIC INDICATORS	
INPUTS	OUTPUTS
Private Consumption	GFCF
Government Consumption	Industrial Production
International Trade (Imports)	International Trade (Exports)

Table 5.1. Country DEA Inputs and Outputs

5.2.2. Sample Data Analysis. Referring to Figure 5.1 and Figure 5.2 let us familiarise ourselves with each country's main economic indicators utilised in our DEA model. Before we begin analysing individual countries' we draw attention to the main economic indicators scale in Figure 5.1 for years 2004 to 2006. In general all the OECD countries consumption and produce has increased substantially over the these three years.

For all three years from 2004 to 2006, Portugal is producing far less than it is consuming compared to the rest of the countries i.e. the outputs Industrial Production and GFCF volumes are below the inputs Private and Government Consumption volumes. Japan, Netherlands, United States and Switzerland are marginally consuming more than they are producing. Interestingly, the Czech Republic, Slovak Republic and Turkey appear to be the star performers, producing far more than they are consuming when compared to the other countries.

Focusing on the level of business activity in the form of GFCF, Australia, Canada, Ireland, Slovak Republic and Turkey are the top five performers. Germany, Poland and Portugal have the lowest GFCF but Poland appears to correct this in 2006. In 2006 countries such as Czech Republic, Hungary, Ireland, Korea, Poland, Slovak Republic and Turkey have the highest Industrial Production that exceeds the rest of the countries by a large gap. In Hungary's case it's Industrial Production has increased and leaped over its private consumption in the three years.

In Figure 5.2 New Zealand, Portugal and Slovak Republic are importing and exporting the least in terms of volume but the United States, Japan and Germany are importing and exporting far more than the rest of the countries. Germany and Japan are exporting more than importing but the gap is not large. The United states, Spain and United Kingdom are noticeably importing more than they are exporting. For the United States, their import volume is massive and the gap between their imports and exports has also widened from 2004 to 2006 while for all other countries it has increased slightly or stayed the same.

The credit ratings for the 25 countries for 2004 to 2007 can be viewed in Table 5.2. The credit ratings given by the OECD contain eight country risk categories. A credit rating of zero represents excellent credit quality, a credit rating of seven represents poor credit quality where the country is highly likely to default on its external debt. All the ratings recorded in Table 5.2 have been taken as of January for each year. Most OECD countries have a credit rating of zero, implying they have excellent credit quality and hardly any risk. The fact that most of the countries in our Table are rated zero makes it difficult to differentiate between these countries.

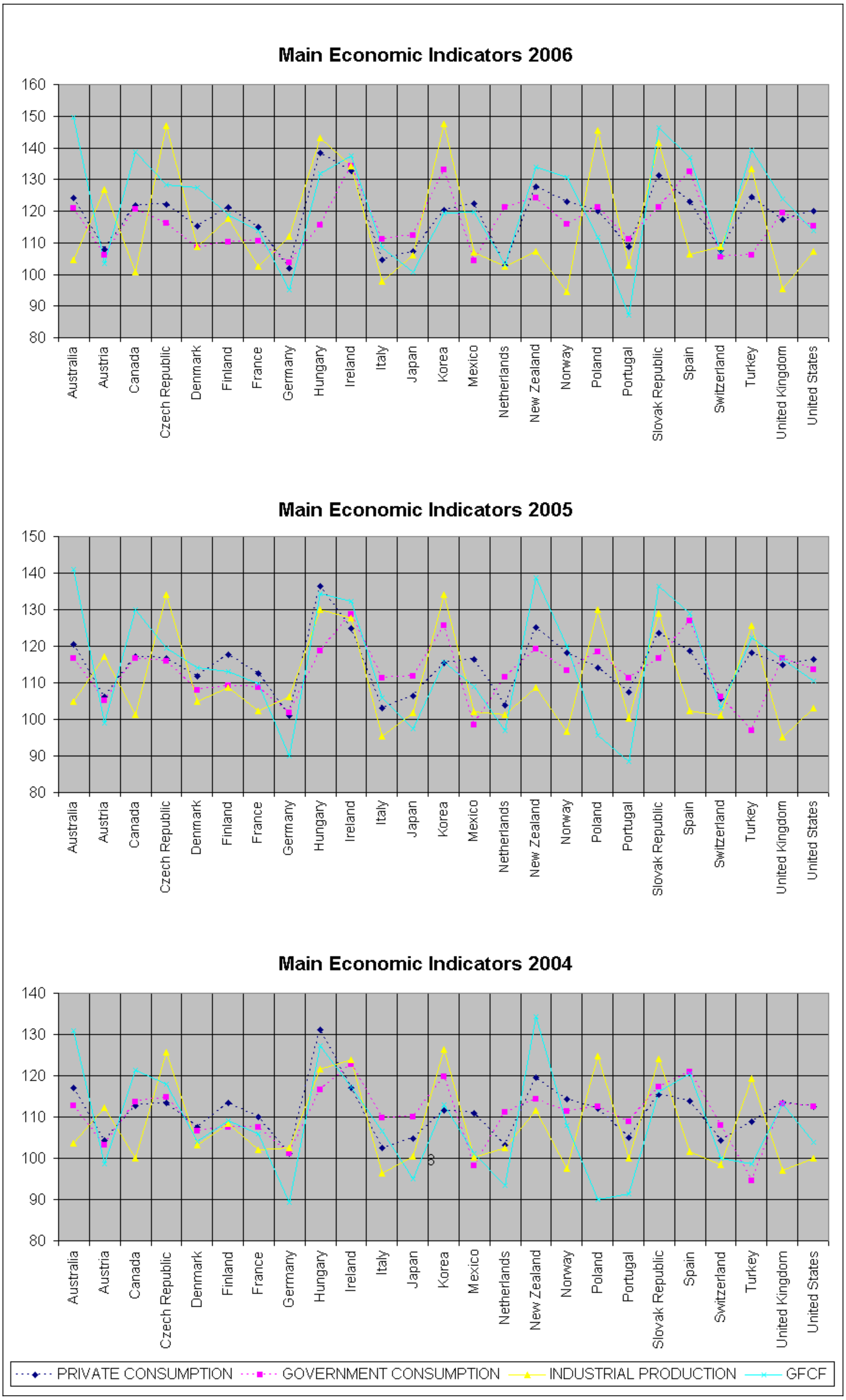

FIGURE 5.1. OECD Main Economic Indicators

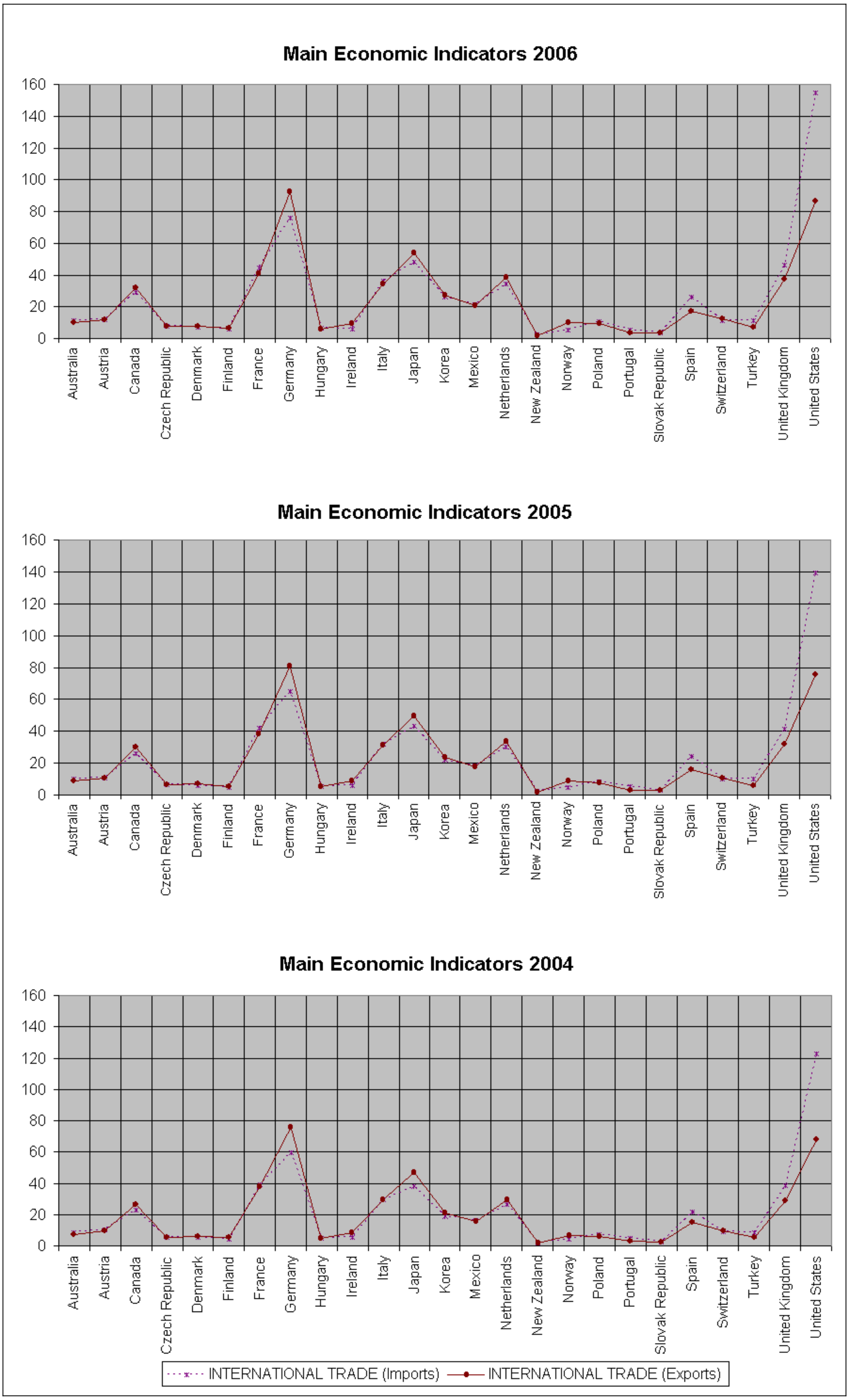

FIGURE 5.2. OECD Main Economic Indicators

In Table 5.2 only six of the 25 countries do not have a rating of zero, they are Czech Republic, Hungary, Mexico, Poland, Slovak Republic and Turkey. Turkey clearly has the worst credit rating in Table 5.2. Turkey is an interesting case, remember that it was one of the countries with the biggest industrial production and business activity which exceeded its private and government consumption. The reason for the poor credit rating is unclear, it could be due to qualitative input such as current policies and possibly political issues. It will be interesting to see what our DEA model produces for this country.

Table 5.2 also clearly show the migration of some of the country ratings over the years 2004 to 2007. The Czeck Republic moved from a credit rating of two to one in 2007. Mexico's rating moved from three to two in 2006. The Slovak Republic country rating appears to be continually climbing, it has improved by two rating categories over the the passed four years. In 2004 and 2005 Slovak Republic has a rating of three, in 2006 its rating was upgraded to a two and in 2007 it's rating was further upgraded to a rating of one. Turkey's credit quality has also improved from a credit rating of six to a credit rating of five in 2005. Admitidly there is little information provided by the credit ratings, but although we only have six countries without a credit rating of zero and of these only four countries migrating between credit ratings we are still interested in comparing the credit ratings with the efficiency results produced by DEA, for who knows what we may find.

The data analysis we have just performed on the economic indicators and the credit ratings has helped us familiarise ourselves with the information provided for each country. With this knowledge we are now in a position to understand and interpret the DEA results.

5.2.3. DEA model application. Now that we have selected and analysed the inputs and outputs for our DEA model, we are ready to apply our DEA model to the 25 Countries. We will use the same DEA model as we did for the corporates. We construct our DEA model to solve several linear programs in equation (5.2.1).

$$
(P'_k)\begin{cases} \max\limits_{z^k\in\mathbb{R}^r, w^k\in\mathbb{R}^s} (z^k, Y^k) \\ s.t. \\ \quad (w^k, X^k) = 1 \\ \quad (z^k, Y^{k'}) \leq (w^k, X^{k'}), \qquad k' = 1, 2, \ldots, N. \\ \quad z^k \geq 0, \quad w^k \geq 0. \end{cases} \tag{5.2.1}
$$

The DEA country model implementation is as follows: We have 25 countries with their corresponding inputs and outputs for years 2004, 2005 and 2006. Run the 25 countries through the DEA model for each year separately. So essentially we have three country data sets, one for 2004, one for 2005 and one for 2006. For each year insert the countries inputs, X, and their outputs, Y, into the DEA model in equation (5.2.1) and solve. The DEA model will calculate the weights for each

Country	OECD Country Credit Ratings			
	Jan-07	Jan-06	Jan-05	Jan-04
Australia	0	0	0	0
Austria	0	0	0	0
Canada	0	0	0	0
Czech Republic	1	2	2	2
Denmark	0	0	0	0
Finland	0	0	0	0
France	0	0	0	0
Germany	0	0	0	0
Hungary	2	2	2	2
Ireland	0	0	0	0
Italy	0	0	0	0
Japan	0	0	0	0
Korea	0	0	0	0
Mexico	2	2	3	3
Netherlands	0	0	0	0
New Zealand	0	0	0	0
Norway	0	0	0	0
Poland	2	2	2	2
Portugal	0	0	0	0
Slovak Republic	1	2	3	3
Spain	0	0	0	0
Switzerland	0	0	0	0
Turkey	5	5	5	6
United Kingdom	0	0	0	0
United States	0	0	0	0

Table 5.2. OECD Country Credit Ratings for 2007-2004

country's inputs and outputs w and z by maximising the weighted sum of outputs over the weighted sum of inputs subject to its constraints. To determine if a country is efficient within a year we calculate the cross-efficiencies using equation (5.2.2),

$$E_{k,k'} = \frac{(u^k, Y^{k'})}{(v^k, X^{k'})}, \quad k' = 1, 2, \ldots, N, \quad k = 1, 2, \ldots, N \tag{5.2.2}$$

subject to the constraints

$$u \geq 0, \quad v \geq 0 \tag{5.2.3}$$

$$E_{k,k'} \leq 1 \quad \text{for all } k' \tag{5.2.4}$$

This simple efficiency $E_{k,k}$ can be used to rank the countries from efficient to least efficient by sorting all $E_{k,k}$'s in descending order. An alternative ranking can be given by the peer efficiency in equation (5.2.5). The peer efficiency can be used to rank the countries by sorting e_k in descending order from most efficient to least efficient, where

$$e_k = \frac{1}{(N-1)} \sum_{k' \neq k} E_{k,k'} \tag{5.2.5}$$

Where too many countries appear efficient in the simple efficiency we may rank order the efficient countries using Θ_k^* in equation (5.2.6) in the CWA, SCWA or MCWA methods. To obtain a ranking of most efficient to least efficient among the already efficient countries we must order Θ_k^* in descending order for any of the three methods CWA, SCWA or MCWA.

$$\Theta_k^* = \frac{\sum_{i=1}^{r} U_i^* y_{i,k}}{\sum_{c=1}^{s} V_c^* x_{c,k}}, \quad k = 1, 2, \ldots, N \tag{5.2.6}$$

Having applied the DEA model we are now ready to analyse the DEA country results.

5.3. Results and Analysis

Table 5.3, 5.4 and 5.5 contains the DEA results for each year. The Tables display the simple efficiency, peer efficiency and maverick results. For the efficient countries the ranking variable Θ_k^* is displayed for the CWA, SCWA and MCWA methods. The country rating for the corresponding year has also been included for easy reference in each Table. For the DEA application to countries, we have selected to rank the countries by the simple efficiency. Where we have several efficient countries we have selected to use the MCWA method to rank these countries. Therefore the first column in the Tables labeled 'Rank' is the ranking of each country where we have sorted firstly by the simple efficiency for all the countries and secondly by the MCWA method for the efficient countries. Where MCWA method gave an equal ranking we sorted these by SCWA method.

Earlier we mentioned that the OECD countries are 'the best' as together they produce two-thirds of the world's goods and services. It is encouraging to see that in all three Tables the countries are either efficient or close to efficient when we focus on the simple efficiency results. Not one country has a simple efficiency below 0.81. This is very different to the results we were seeing for the corporate sub-segments where corporates received simple efficiency scores within the full range from one to zero. These results are also encouraging as they confirm that these countries form a homogeneous set and no 'single country' is operating vastly differently to the others. The peer efficiency and maverick indicators also support this view of the countries being a homogeneous set. Peer efficiency in this context refers to how all the countries view a particular country without taking its own efficiency into account. Notice that all the peer efficiencies are in a tight range between 0.5 and

Rank	Country	OECD Rating Rating	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)	CWA (Θ_k^{C*})	SCWA (Θ_k^{S*})	MCWA (Θ_k^{M*})
1	Slovak Republic	2	1	0.909	0.100	1.000	1.101	1.045
2	Norway	0	1	0.788	0.270	0.968	1.017	1.022
3	Turkey	5	1	0.806	0.241	1.000	1.121	1.022
4	Australia	0	1	0.769	0.300	0.977	1.053	1.008
5	Czech Republic	2	1	0.851	0.175	0.963	1.071	1.000
6	Ireland	0	1	0.820	0.220	0.953	1.049	1.000
7	Germany	0	1	0.732	0.367	1.000	1.000	1.000
8	Hungary	2	1	0.805	0.243	0.916	1.000	0.957
9	New Zealand	0	1	0.823	0.214	0.911	1.004	0.953
10	Korea	0	1	0.737	0.358	0.912	1.027	0.931
11	Poland	2	1	0.760	0.316	0.872	1.000	0.901
12	United States	0	0.999	0.556	0.797			
13	Canada	0	0.995	0.715	0.392			
14	Denmark	0	0.981	0.801	0.224			
15	Austria	0	0.974	0.762	0.278			
16	Spain	0	0.954	0.651	0.466			
17	Netherlands	0	0.950	0.677	0.403			
18	Italy	0	0.947	0.663	0.428			
19	Mexico	2	0.944	0.690	0.369			
20	United Kingdom	0	0.942	0.626	0.504			
21	Finland	0	0.942	0.784	0.203			
22	Japan	0	0.922	0.673	0.370			
23	Switzerland	0	0.917	0.737	0.244			
24	France	0	0.914	0.648	0.410			
25	Portugal	0	0.816	0.649	0.258			

TABLE 5.3. DEA Efficiencies for OECD countries in 2006

0.93. As a bi-product of the tight ranges of the simple efficiency and peer efficiency, we do not have any mavericks. Not one country has a maverick result over one, which is quite low when we compare with the maverick results seen in the corporate application.

In 2006 we have eleven efficient countries with a simple efficiency equal to one, of which the Slovak Republic is ranked number one. In 2005 we also have eleven efficient countries but Canada and the United States appear efficient instead of Poland and Hungary. In 2005 Turkey is ranked as the most efficient country followed by the Slovak Republic. In 2004 again we have eleven efficient countries with Poland replacing the United States. New Zealand is ranked the most efficient country in 2004 with the Slovak Republic and Turkey in third and fourth ranking. As surprising as these results may be, when we refer back to the sample data analysis it is clear why our DEA model is ranking in this way. Let us investigate the number one ranked countries a little deeper. For 2006, scan across all the countries in Figure 5.1 and then focus on the Slovak Republic. Notice that not only are its output indicators, GFCF and Industrial production, one of the highest (in the case

Rank	Country	OECD Rating Rating	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)	CWA (Θ_k^{C*})	SCWA (Θ_k^{S*})	MCWA (Θ_k^{M*})
1	Turkey	5	1	0.802	0.247	1.000	1.090	1.066
2	Slovak Republic	3	1	0.922	0.084	0.986	1.039	1.036
3	Australia	0	1	0.780	0.282	1.000	1.059	1.028
4	Canada	0	1	0.757	0.321	0.994	1.121	1.009
5	Germany	0	1	0.791	0.264	1.000	1.304	1.000
6	Ireland	0	1	0.864	0.158	0.954	1.021	1.000
7	New Zealand	0	1	0.872	0.147	0.961	1.002	1.000
8	Czech Republic	2	1	0.854	0.171	0.936	1.000	0.987
9	Korea	0	1	0.784	0.276	0.938	1.049	0.977
10	Norway	0	1	0.807	0.240	0.927	1.000	0.970
11	United States	0	1	0.574	0.742	0.820	1.000	0.801
12	Poland	2	0.990	0.743	0.333			
13	Hungary	2	0.972	0.808	0.203			
14	Austria	0	0.969	0.770	0.260			
15	Italy	0	0.964	0.703	0.371			
16	Spain	0	0.955	0.664	0.438			
17	Mexico	3	0.938	0.705	0.331			
18	France	0	0.938	0.688	0.364			
19	Denmark	0	0.937	0.791	0.184			
20	Japan	0	0.932	0.725	0.286			
21	United Kingdom	0	0.930	0.643	0.445			
22	Finland	0	0.924	0.784	0.178			
23	Netherlands	0	0.923	0.718	0.287			
24	Switzerland	0	0.907	0.744	0.220			
25	Portugal	0	0.831	0.677	0.228			

TABLE 5.4. DEA Efficiencies for OECD countries in 2005

of GFCF it is the highest), but its inputs Private and Government Consumption are relatively low in comparison. In Figure 5.1 for 2005, Turkey is producing more output than consuming input and notice the incredibly low input of Government Consumption. This is implying Turkey is producing much more output with less input compared to the other countries and the DEA model and MCWA method is identifying this immediately and placing Turkey in the number one rank spot for efficient countries. In 2004 New Zealand had the highest GFCF with Government and Private Consumption average in volume, but we can see it did not continue growing compared to the rest of the countries and fell off the top end of the simple efficiency Tables in 2005 and 2006.

In all three years 2004 to 2006, Portugal is ranked as the least efficient country in Table 5.3, 5.4 and 5.5. Going back to our sample data analysis in Figure 5.1 we observed that Portugal was consuming more that it was producing for all three years so this result makes intuitive sense based on the data.

Rank	Country	OECD Rating Rating	Simple Efficiency ($E_{k,k}$)	Peer Efficiency (e_k)	Maverick (M_k)	CWA (Θ_k^{C*})	SCWA (Θ_k^{S*})	MCWA (Θ_k^{M*})
1	New Zealand	0	1	0.929	0.077	1.000	1.185	1.042
2	Ireland	0	1	0.877	0.141	0.990	1.123	1.026
3	Australia	0	1	0.787	0.270	0.989	1.151	1.009
4	Turkey	6	1	0.798	0.254	1.000	1.148	1.000
5	Slovak Republic	3	1	0.900	0.111	0.975	1.124	1.000
6	Canada	0	1	0.779	0.283	0.987	1.094	1.000
7	Korea	0	1	0.811	0.234	1.000	1.093	1.000
8	Germany	0	1	0.818	0.223	1.000	1.000	1.000
9	Czech Republic	2	1	0.861	0.161	0.970	1.115	0.992
10	Norway	0	1	0.804	0.244	0.936	1.080	0.979
11	Poland	2	1	0.747	0.338	0.895	1.000	0.896
12	Italy	0	0.995	0.736	0.352			
13	Hungary	2	0.990	0.815	0.214			
14	United States	0	0.988	0.580	0.703			
15	Finland	0	0.973	0.825	0.180			
16	Austria	0	0.971	0.791	0.228			
17	Spain	0	0.968	0.687	0.410			
18	Japan	0	0.965	0.767	0.259			
19	Mexico	3	0.956	0.734	0.303			
20	France	0	0.953	0.723	0.317			
21	United Kingdom	0	0.944	0.665	0.420			
22	Denmark	0	0.934	0.798	0.170			
23	Netherlands	0	0.917	0.734	0.249			
24	Switzerland	0	0.910	0.750	0.213			
25	Portugal	0	0.865	0.717	0.207			

TABLE 5.5. DEA Efficiencies for OECD countries in 2004

When we analyse the DEA rankings alongside the OECD rating we notice that in general the majority of the poorly rated countries are ranked as efficient or very close to efficient. In fact, of the six countries not rated zero only Mexico never appears as efficient on any of the Tables. In 2006 five of the six countries not rated zero are efficient. In 2005 there are three efficient countries with ratings other than zero, with another two countries Poland and Hungary ranked just under the efficient countries. In 2004 four of the six countries not rated zero are efficient with Hungary ranked just under the efficient countries. These country results are clearly showing us that rating and DEA ranking do not appear to agree and are measuring something the other is not.

When we compare the efficiency results with the credit rating we also observe something else which is very interesting. Notice that of the four countries (Czech Republic, Mexico, Slovak Republic and Turkey) who's rating has migrated to a better credit rating over the years 2004 to 2007 as seen in Table 5.2, three of them were identified in our DEA model as operating efficiently. The Slovak Republic was classified as an efficient country by our DEA model and ranked as the third and fourth efficient country in 2004 and 2005 respectively when it had a credit rating of three. In 2006 its credit rating improved to a credit rating of two. In 2006 our DEA model ranked the Slovak Republic as the most efficient country and in 2007 its credit rating improved once more to a credit rating of one. Could this be coincidence, or could it be that with the data we provided our DEA model, it correctly identified the growth and superior performance that the Slovak Republic was making? Could it be that our DEA model identified that the Slovak Republic was indeed performing efficiently and due to its performance it's credit rating improved? We see a similar case for the Czech Republic and Turkey. Both these countries have been identified as efficient by our DEA model and both have credit ratings that have improved.

5.4. Conclusion

In this chapter we applied DEA to OECD countries using thoroughly analysed, reliable and comparative OECD economic data. We naturally expected to see countries in the G10 (Canada, France, Germany, Italy, Japan, Netherlands, Spain, Switzerland, United Kingdom and United States) that have the best credit ratings ranked at the top of the efficiency Tables. Instead very few G10 countries were identified as efficient in our DEA model. Countries such as the Slovak Republic, Turkey, and Czech Republic that did not have the best ratings were identified as efficient. This result clearly indicates that simple efficiency produced by our DEA model and credit rating are measuring something different. The credit rating is measuring the countries likelihood to default on its external debt. The credit rating is made up of quantitative information such as economic indicators as well as qualitative information such as the countries policies and political stability. The simple efficiency is measuring if a country is operating efficiently based on it's economic indicators, by what it is producing (output) and consuming (input) relative to the other countries. Policy and political issues are not taken into account. Instead efficiency is supplying us with a condensed view of a country's economic indicators and how successful every country is, in comparison to other countries and their resources.

Based on our data sample we clearly observed that countries that produced more output with less input relative to rest of the countries were indeed found to be operating more efficiently, so our DEA results are valid.

We observed that countries with poor credit ratings and excellent efficiency scores migrated to better credit ratings. This is certainly not a coincidence. We noticed in the case of the Slovak Republic, when it was ranked the most efficient in 2006, the credit rating improved the very next year in 2007. A possible explanation based on the make up of the efficiency and the credit rating is as follows: A country's credit rating can be improved for various reasons for example due to new accepted policies and standards being adopted, political situations stabilising or the general produce a country generates. It is on the last point where credit rating and efficiency align. A country will only be identified as efficient if it is operating efficiently as shown by its main economic indicators. For the Slovak Republic, Czech Republic and Turkey this appears to be the case. For Mexico, ranked inefficient, the credit rating improved over the years but our DEA model clearly identified it was not due to it's economic indicators improving. Our DEA model is clearly indicating that a country's credit rating has the potential to improve if it is efficient.

DEA efficiency is identifying the opportunity and potential a country has, based on its economic indicators. Efficiency appears to have nothing to do with 'risk' in a qualitative dimension but everything to do with risk in a quantitative dimension when we look at the pure economics of a country. The credit rating appears to dilute the quantitative aspects of a country and therefore would miss out on any business opportunities to spot an acceleration in growth and wealth in a country. In this context, DEA has a clear function as a business tool and a complement to the credit rating. It would be in a credit lender's best interest to lend to countries with poor credit ratings and high efficiency scores. As a credit lender we could charge more for credit because the current credit rating is poor. We would be confident that the country is operating efficiently based on DEA results. The DEA results also indicate that the efficient country has the potential for credit risk to improve. So, finally we would reap the benefits of an accelerating economy at high profit because we engaged in credit lending early. If used in this way the DEA model could be an excellent strategic business tool which complements the rating and its forecasts for the future.

In summary efficiency is identifying those countries that are performing better than the rest based on economic indicators. Efficiency provides a good differentiation tool between countries. Efficiency and credit rating are measuring two different but both important items. Efficiency indicates if a country economy is accelerating compared to the rest and has the potential to identify future improvements in credit rating based on country performance. With the clear results presented above DEA should be used as a strategic business tool and a complement to the credit rating tools. If reliable and comparative country information becomes available for non-member OECD countries that also includes samples of country ratings that have deteriorated over time, future research could entail investigating if a similar relationship holds true for non-member OECD countries.

6

Conclusions

Having examined DEA, its extensions and applications we have found it to be a suitable tool for credit risk measurement. DEA compares a population of observations against a 'best practice' observation. It creates an envelopment frontier based on the populations' relative productive efficiency. This frontier is a surface with invaluable information using multiple inputs and multiple outputs for the comparison between peers. It is this unique quality that separates DEA from traditional methods such as regression analysis and enables DEA to detect relationships that remain concealed from other methods. In the field of mathematics this non-parametric approach is well established and in the field of economics it is well supported.

In a relevant literature study we discussed the comparison of US banks' CAMEL ratings to DEA efficiency and found that more efficient banks tend to be higher performers and safer institutions. In our application to Corporate Credit Risk we investigated if there was a similar relationship between corporate credit ratings and DEA efficiency. We found that although we faced challenges with the corporate credit rating data clustered in two credit ratings, there is such a relationship. Corporates receiving better credit ratings are generally more efficient. We demonstrated the ease and benefit of applying DEA to differentiate between several corporates within a single credit rating grade. With these findings under the application of DEA to corporate entities, we recommend DEA be used as a supplementary tool to corporate credit rating models to facilitate corporate differentiation within clustered credit rating grades.

In the application of DEA to Country Credit Risk we found that efficiency identified those countries that are star performers based on their economic indicators. Similar to corporates, efficiency provides a good differentiation tool between countries within a single credit rating grade. We observed the differences between the efficiency of a country and its credit rating. These measures did not have any relationship at first glance. Country ratings were measuring the likelihood of a country servicing its external debt that included quantitative economic indicators and qualitative (political and policy) elements. Efficiency only focused on the quantitative economic indicators. Later when we investigated the relationship of efficiency and credit rating through time we found that efficiency indicates if a country's economy is accelerating compared to the rest. Efficiency showed the potential to identify future improvements in credit ratings based on country performance and economic indicators. With the findings under the application of DEA to countries, we highly recommend the use of DEA as a strategic business tool used in guiding businesses to high yields when investing in countries. In addition DEA can be engaged as a complement to the credit rating tools used to further differentiate and identify credit rating migrations.

A study undertaken by [**8**] tested efficiency results produced by DEA models through time as an alternative or supplement to the Basel concept known as 'risk adjusted capital'. In the study DEA was able to detect a worsening banking situation over time which later played out [**8**]. Although the study was undertaken for illustrative purposes it vindicates using DEA as a risk measurement tool. Therefore, future research should be undertaken in the corporate lending sector, specifically around comparing credit ratings and efficiency through time. This should be done on a data sample where credit ratings are more evenly distributed across the credit rating scale to identify if the same common sense relationships still exist. Beyond this dissertation valuable research should also be undertaken on the application of DEA on non-OECD member countries if and when reliable and comparative country information becomes available.

In conclusion, improved management of credit risk by banks can be greatly enhanced by the implementation of DEA. In competitive markets only the strong survive. To survive we need to achieve high levels of performance through continued improvements and learning. DEA and its extensions provide insight into competitive markets drawing on comparisons between peers. The use of DEA will improve credit risk measurement and management.

Glossary

AIRB: Advanced Internal Ratings Based Approach

Asset Classes: Are a Basel II categorisation of exposures namely corporate (corporates are further divided into large corporates, small-to-medium enterprises (SME), project finance (PF), object finance (OF), commodity finance (CF), Income-producing real estate (IPRE), High-volatility commercial real estate(HVCRE), sovereigns, banks, equities and retail.

Backtesting: Comparing estimates to actual observed results.

Basel II Accord: An international agreement to establish capital adequacy standards for banks.

Benchmarking: To take measurement against a reference point.

CAMEL: Capital adequacy, Asset quality, Management quality, Earnings ability, Liquidity.

CAMEL Ratings: CAMEL ratings are a common measure used by regulators for evaluating a bank's financial health. The use of CAMEL ratings is wide-spread among regulators. The evaluation factors are as follows: Capital adequacy, Asset quality, Management quality, Earnings ability, Liquidity. Each of the factors are scored 1 to 5 where 1 is the strongest rating. An overall composite CAMEL rating, also ranging from one to five, is then developed from this evaluation. Each rating level has a definition.

Cleaning Data: The process whereby raw data is analysed and filtered from missing values and outliers.

Constraints: A set of equalities or inequalities that must be satisfied.

Credit analyst: An individual that has extensive credit risk knowledge involved in assessing the credit risk of entities.

Credit Expert: see Credit analyst.

Credit Manager: see Credit analyst.

Credit Rating: An assessment of credit risk usually given on categorical risk scale.

Credit Rating Grade: see Credit rating.

Credit Rating Migration: The movement of an entity's credit rating from one credit rating grade to another.

Credit Rating Model: Models used to produce credit risk ratings.

Credit Risk: Exposure to the risk of a loss arising from default or credit downgrade.

CWA: Common Weights Analysis. A method used to rank efficient DMUs in a produced by a DEA model.

Default Grade: A credit rating grade that represents default.

DMU: Decision Making Units. A collection of entities evaluated in a DEA model that consume several inputs and produce several outputs.

EAD: Exposure at Default. The amount an institution is exposed to when an entity defaults on its debt.

EBITDA: An indicator of a companies financial performance used to analyse and compare profitability between companies and industries because it eliminate the effects of financing and counting decisions.

Economic Efficiency: An measure that focuses on and entities input and output data as well as market price data.

Efficiency: A measure between zero and unity, for DEA purposes defined as the weighted sum of inputs divided by the weighted sum of outputs. See also economic efficiency and productive efficiency.

Efficient Frontier: A non-parametric piece-wise surface constructed by a DEA model by efficient DMUs.

EL: Expected Loss. The amount an institution expects to loose. Expected loss is calculated by multiplying PD, LGD and EAD.

Envelopment Surface: See efficient frontier.

FIRB: Foundation Internal Ratings Based Approach.

GDP: Gross Domestic Product is one of the ways for measuring the size of a countries economy.

GFCF: Gross Fixed Capital Formation. Measures a countries level of business activity.

Homogeneous group: A group with similar characteristics and risks.

IIP: Index of Industrial Production

Internal Credit Rating: credit ratings assigned by an institution internally that are usually not disclosed to the public.

IRB: Internal Ratings Based Approach

JSE: Johannesburg Securities Exchange

LGD: Loss Given Default. The loss an institution will incur given an entity defaults.

Linear Programming: An optimization problem where the constraints and the objective function are all linear.

Listed Corporates: Corporate entities that are listed on the Johannesburg Securities Exchange.

MCWA: Most Compromising Set of Indices' Weights Analysis. A recent method used to rank efficient DMUs produced by a DEA model.

Non-Archimedean Quantity: A variable with a positive value very close to zero and is smaller than *any* positive real number.

Non-Linear Programming: An optimization problem where the constraints and or the objective function can be a non-linear function of the decision variables.

Non-Parametric Models: A distribution free model. Non-parametric models are best described when compared to parametric models. They differ from parametric models in that the model structure is not specified a priori, but is instead determined from data. The term nonparametric is not meant to imply that such models completely lack parameters; rather, the number and nature of the parameters is flexible and not fixed in advance.

Objective Function: A function in linear or nonlinear program that is either being minimized or maximized.
OECD: Organisation for Economic Co-operation and Development
OEEC: Organisation for European Economic Co-operation
OLS: Ordinary least-square method
Optimal Solution: A vector x which is both feasible (satisfying the constraints) and optimal (obtaining the largest or smallest objective value).
Overrides: Where human intervention plays a role in the final outcome of a credit rating within a credit rating model.

Parametric: A function is 'parametric' in a given context if its functional form is known to the economist.
Parametric Approach: Any approach computed by procedures that assume the data were drawn from a particular distribution.
Pareto dominance: See Pareto efficient.
Pareto Efficiency: Given a set of alternative allocations and a set of individuals, a movement from one alternative allocation to another that can make at least one individual better off, without making any other individual worse off is called a Pareto efficient or Pareto optimisation. An allocation of resources is Pareto efficient or Pareto optimal when no further Pareto improvements can be made. Within a DEA context pareto efficient means that a DMU is fully efficient if and only if it is not possible to improve any input or output without worsening some other input or output.
Pareto-Koopmans efficiency: See Pareto efficient.
PD: Probability of Default. The likelihood of an entity defaulting within one year.
Peer Group Analysis: A detailed comparison of a set of homogeneous entities.
Productive Efficiency: A measure that focuses on levels of inputs relative to levels of outputs.

Recession: A temporary decline in economic activity.
Regression: The relationship between selected values of x and observed values of y from which the most probable value of y can be predicted for any given value of x.
Regression Analysis: The use of regression to make quantitative predictions of one variable from the values of another.
Risk Adjusted Capital: Stems from the Basel agreement for dealing with adjusting the net worth requirements of individual banks in order to allow for risk differences arising from the various components of each bank's portfolio of loans and investments.

SCWA: Super Common Weight Analysis. A method used to rank efficient DMUs produced by a DEA model.

SME: Small to medium Enterprises. Small corporate entities.

Unlisted Corporates: Corporates privately owned and not listed on the Johannesburg Securities Exchange.

VAR: Value at Risk. A loss that will not be exceeded at some specified confidence level.

References

[1] Angulo-Meza, L. and Lins, M.P.E. *Review of Methods for Increasing Discrimination in Data Envelopment Analysis*, Annals of Operations Research 116, 225-242, Kluwer Academic Publishers (2002).

[2] Basel Committee on Banking Supervision, *International Convergence of Capital Measurement and Capital Standards*, Bank for International Settlements, Basel, Switzerland, July (2006),

[3] Bauer, P.W., Berger, A.N., Ferrier, G.D. and Humphrey, D.B. *Consistency Conditions for Regulatory Analysis of Financial Institutions: A comparison of Frontier Efficiency Methods,*, Journal of Economics and Business 50 (March/April):85-114 (1998).

[4] Charnes, A., Cooper, W.W., Lewin, A.Y. and Seiford, L.M., *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston (1994).

[5] Charnes, A., Cooper, W.W. and Rhodes, E. *Measuring the efficiency of decision making units*, European Journal of Operational Research 2 (1978) 429-444.

[6] Charnes, A., Cooper, W.W., Sears, M. and Zlobec, S. *Efficiency Evaluations in Perturbed Data Envelopment Analysis*, Presented at the Second International Conference on Parametric Optimisation, Eisenach, Germany, June 29-30, (1989).

[7] Charnes, A. and Neralic, L. *Sensitivity Analysis of the Additive Model in Data Envelopment Analysis*, Center for Cybernetic Studies, Research Report CCS 628, University of Texas, Austin, May (1989).

[8] Cooper, W.W., Seiford, L.M. and Tone, K. *Introduction to Data Envelopment Analysis and its Uses*, Springer, United States, 2006.

[9] Doyle, J. and Green, R. *Efficiency and Cross-Efficiency in DEA: Derivations, Meanings and Uses*, The Journal of Operational Research Society, Vol. 45, No.5. (May 1994) 567-578.

[10] Färe, R.S., Grosskopf, and Lovell, C.A.K. *Production Frontiers*, Cambridge: Cambridge University Press (1994).

[11] Liu, F.H., Peng, H.H. *DEA and common weight analysis for ranking DMUs*, Computers and Operations Research, (2004).

[12] Liu, F.H., Peng, H.H. and Chang, H.W. *Ranking DEA Efficient Units with the Most Compromising Common Weights*, The Sixth International Symposium on Operational Research and Its Applications (ISORA '06), Xinjiang, China, August (2006) 219-234.

[13] OECD http://www.oecd.org

[14] OECD *OECD Main Economic Indicators Explanatory Notes*, http://www.oecd.org/dataoecd (2007).

[15] OECD *OECD Main Economic Indicators*, http://www.oecd.org/dataoecd/ April (2007).

[16] Siems, T.F. and Barr, R.S. *Benchmarking the Productive Efficiency of U.S. Banks*, Federal Reserve Bank of Dallas, Financial Industry Studies, December (1998) 11-24.

[17] South African Reserve Bank (SARB), *Press Release by the Registrar of banks on the 2004 annual report of the bank supervision department*, http://www.reservebank.co.za, 29 June (2005).

[18] The Star, *Banks get heads-up on Basel II compliance*, Business Report, South Africa's national financial daily, 26 February (2007).

[19] Talluri, S. *Data Envelopment Analysis: Models and Extensions*, Decision Line (2000). Federal Reserve Bank of St Louis (1992).

[20] Yue, P. *Data envelopment analysis and commercial bank performance: a primer with applications to Missouri banks*, Federal Reserve Bank of St Louis (1992).