Incorporating Ratios in DEA—Applications to Real Data

written by

Sanaz Sigaroudi

Thesis advisor:
Dr. Joseph C. Paradi

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Mechanical and Industrial Engineering
University of Toronto

Copyright © 2009 by Sanaz Sigaroudi
Abstract

Incorporating Ratios in DEA—Applications to Real Data

Sanaz Sigaroudi

Master of Applied Science

Graduate Department of Mechanical and Industrial Engineering

University of Toronto

2009

In the standard Data Envelopment Analysis (DEA), the strong disposability and convexity axioms along with the variable/constant return to scale assumption provide a good estimation of the production possibility set and the efficient frontier. However, when data contains some or all measures represented by ratios, the standard DEA fails to generate an accurate efficient frontier. This problem has been addressed by a number of researchers and models have been proposed to solve the problem. This thesis proposes a “Maximized Slack Model” as a second stage to an existing model. This work implements a two phase modified model in MATLAB (since no existing DEA software can handle ratios) and with this new tool, compares the results of our proposed model against the results from two other standard DEA models for a real example with ratio and non-ratio measures.

Then we propose different approaches to get a close approximation of the convex hull of the production possibility set as well as the frontier when ratio variables are present on the side of the desired orientation.
Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Joseph C. Paradi for giving me the opportunity to be part of his team at CMTE. I am indebted to him for his life lessons, continuous support and patience. His passion for having an active social and scientific life, even after retirement, has always been a source of inspiration to me.

I am grateful to my committee members Professor Y. Lawryshyn, Professor R. Kwon and Professor B. Balcioglu for providing me with constructive comments and insightful feedback. My appreciation goes to Dr. Emrouznejad for providing access to his paper and the proposed model therein, which became a motivation for this work.

During my time at CMTE I learned a lot from discussions on DEA and related mathematical modeling with my colleagues. In particular I am thankful to Dr. Barak Edelstein for teaching me the DEA concept and Dr. Haiyan Zu for answering my questions and sharing her expertise with me. Dalia Sheriff’s knowledge on ratios and Leili Javanmardi knowledge on MATLAB were also very helpful at the early stages of this work.

I owe an immeasurable debt of gratitude to Dr. Judy Farvolden for her warm welcome when I first arrived in Canada and her ongoing scientific and non-scientific support and advice throughout my research career. My defense presentation would not be successful without Judy and Haiyan’s help.

My friends at CMTE Angela Tran, Pulkit Gupta, Laleh Kobari, Sukrit Ganguly, Steve Frensch, Susuan Mohammadzadeh, Erin Kim, Colin Powel, Joe Pun, and Justin Toupin have made my student life a joyful experience. I also thank Sau-Yan Lee for his help whenever I had a problem with computer and network. My appreciation goes to Professor Balcioglu for trusting me and giving me the opportunity to gain invaluable teaching experience. I would also like to thank the staff at the MIE graduate office, mostly Brenda Fung, Donna Liu, and Joe Nazal for helping me out through the administrative parts of the process.
I have been always blessed with amazingly nice friends. I am especially grateful to Ghazal and Niayesh for being the best neighbors and friends one can ever ask for. My kind relatives, Makhmal and Anoosh, and their kids, Maryam and Mona, have always welcomed me warmly into their house in Toronto at many late nights when commuting to Waterloo was hard. Yaser and Faranak, my cousin Solmaz, Iman and Golnaz, Sara and Ali, as well as Afsaneh and Reza have made the life experience in Waterloo and Toronto, far from my home country, Iran, a wonderful, adventurous and joyful one. I like to say “merci” to my dearest friend, Mahnaz and her family, for being so kind to me and my husband no matter what we do or where we go.

I am indebted to my parents, Habibeh Ghaznavi and Dr. Issa Sigaroudi, and my brother, Sina, for loving me and believing in me and giving me the full freedom to explore the world of unknowns in my own way. I am also very lucky to have caring and loving parents-in-law, Zahra Alemzadeh and Dr. Asadollah Razavi who keep me in their good prayers. My sisters-in-law Maryam, Razieh, and Marziyeh and their families have been always there to cheer for me and support me. I would also like to thank my aunt, Alieh, my cousin Mahnaz, and my husband’s uncle, M. Ali, and his family for their kindness and support during our stay in the United States and Canada.

I would not have been able to continue my studies without the love, support, assistance and encouragement of my beloved husband, Mohsen Razavi, throughout theses years. His love and kind heart are the most precious gifts I have ever got in this world. This thesis is dedicated to him.
Contents

1 Introduction 1
  1.1 Introduction ......................................................... 1

2 RA and DEA 4
  2.1 RA ................................................................. 4
    2.1.1 How RA is measured? .......................................... 4
    2.1.2 RA and linear programming representation .................. 5
  2.2 Data Envelopment Analysis ......................................... 7
    2.2.1 Introduction .................................................... 7
    2.2.2 DEA Basic assumptions and Models .......................... 8
    2.2.3 Production Possibility Set ................................... 8
    2.2.4 Frontier ........................................................ 9
    2.2.5 Efficiency Definition ......................................... 10
    2.2.6 Orientation ..................................................... 12
    2.2.7 Additive model ................................................ 12
  2.3 Linkage between Data Envelopment Analysis and Ratio Analysis 12
    2.3.1 Comparing DEA and RA ....................................... 12
    2.3.2 Combining DEA and RA ....................................... 14
  2.4 Advantages of incorporating ratios in DEA ........................ 16
3 Frontier Assessment

3.1 Deterministic Frontier Estimators .............................................. 18
  3.1.1 Free Disposal Hull Frontier ................................................. 18
  3.1.2 Various Returns to Scale Frontiers ...................................... 18
  3.1.3 Constant Returns to Scale Frontier .................................... 19

3.2 Probabilistic Frontier Estimators ............................................ 19
  3.2.1 Partial m Frontier .......................................................... 20
  3.2.2 Quantile Frontier .......................................................... 22
  3.2.3 Practical Frontiers ......................................................... 22

4 Incorporating ratios in DEA ....................................................... 24

4.1 Modified DEA models ............................................................ 24
  4.1.1 Model I ................................................................. 25
  4.1.2 Model II ................................................................. 27

4.2 Motivating Example ............................................................... 30

4.3 Nonlinearity and Orientation .................................................. 38

5 Comparing the Conventional and the Modified DEA Solver ............ 41

5.1 Use of DEA in Banking Industry .............................................. 41

5.2 Evaluation Criteria ............................................................... 43
  5.2.1 Profitability model ......................................................... 44

5.3 Data and summary results ....................................................... 46
  5.3.1 Correlation ................................................................. 47
  5.3.2 “Non-ratio” Profitability Model, DEA solver results ................ 48
  5.3.3 “Conventional Ratio” Profitability Model, DEA solver results .... 50
  5.3.4 “Modified Ratio” Profitability Model, our MATLAB coded DEA
       solver results .............................................................. 51

5.4 Comparing the Efficiency Scores ............................................. 52
# List of Tables

4.1 Lasik Equipment Information ........................................... 31  
4.2 Comprehensive Lasik Equipment Information ....................... 31  
4.3 Hypothetical DMU Information ........................................... 32  

5.1 Choice of input and output for the profitability model ............. 45  
5.2 Summary of the Input correlations ..................................... 47  
5.3 Summary of the Output correlations in non-ratio and modified model . 49  
5.4 Locally compared .......................................................... 55  
5.5 Input and Output Projections for DMU 58. Note: ROR = Rate of Return 56  
5.6 Distance between DMU 58 and the target ............................... 58  
5.7 Detailed road map to achieve ratio improvements for DMU 58 .......... 59  

6.1 Areas differentiating the top ten efficient branches from average .... 66  
6.2 Sensitivity Analysis ....................................................... 69
List of Figures

2.1 DEA basics ................................................................. 10

3.1 Shapes of conventional frontiers:FDH, BCC, and CCR ................. 19

4.1 Shapes of projections of conventional(red) and modified(black) frontiers on the output plane. ......................................................... 33

4.2 Shapes of conventional(red) and modified(black) 2D frontiers ......... 34

4.3 Feasible DMUs generated by conventional(red) and modified(black) PPS estimators ................................................................. 35

4.4 PPS boundaries in 3D: conventional (red), modified (black) .......... 36

4.5 Forming the efficient frontier, output oriented. DEA conventional(red), DEA modified(black). ......................................................... 36

4.6 Forming the efficient frontier, output oriented. DEA conventional(red), DEA modified(black). ......................................................... 37

4.7 Frontier facets, input orientation. ............................................ 37

5.1 Nation wide comparison of scores ........................................... 54

6.1 Average Input Levels of Regions Branches .............................. 63

6.2 Average Output Levels of Regions Branches .............................. 64

6.3 Average expenses of top ten branches .................................... 66

6.4 Average rate of returns of top ten branches .............................. 67

6.5 Correlation between Modified scores and Bank’s Internal Metrics .... 70
Chapter 1

Introduction

1.1 Introduction

Data Envelopment Analysis (DEA) and Ratio Analysis (RA) are two methods for benchmarking production units’ productivity, profitability or any other criteria that could be assessed based on the available input and output variables. Although the two methods share the same goal, their approach is different in many aspects.

In brief, DEA is a nonparametric method that generates a single relative efficiency score, while considering multiple inputs and outputs simultaneously. RA, in contrast, uses the ratio of a single output to a single input and generates a relative efficiency score by dividing the aforesaid ratio by the corresponding “best performer’s” ratio, on this specific ratio definition. DEA and individual ratios agree weakly on the performance of the unit and it is because DEA reflects the overall efficiency based on a “ratio” of the weighted sum of relevant outputs over inputs and the RA is a factor specific efficiency. Despite DEA’s popularity among academics, it is not widely used as a practical tool for performance assessment in industry. In contrast, RA has been extensively used, despite its well documented limitations and biases [Kriv 08].

It seems that the major hurdle in DEA deployment is its different language in com-
municating the results to management. A language that involves RA in explaining DEA results would make it more understandable to management and, as a result, more appealing to industry. This new language requires the use of desired ratios as the inputs and/or outputs of the decision making units (DMUs).

In 2003, Hollingsworth et al. [Holl 03], for the first time, pointed out the problem of using ratios as inputs or outputs in the CCR model. Later on in 2008, Emrouznejad et al [Emro 09] highlighted the convexity axiom violation when using ratios in inputs or outputs in the BCC model and proposed a modified model when ratios are present.

The purpose of this thesis is to a) study the problems in conventional DEA models in estimating the Production Possibility Set (PPS) and the frontier when we have ratios as variables, b) propose a “Maximized Slack” model and build a two phase DEA model, c) develop a MATLAB code for the suggested model and apply the model to real banking data, evaluate the possible improvements and savings, and report the suggestions in a more familiar language to management, d) compare the results to two other DEA models, and e) deal with the nonlinearity of the suggested model when ratios are on the side of the model orientation.

This thesis is structured as follows:

- Chapter 2 provides an overview of RA and DEA theory and compares the use of the two in different cases.

- Chapter 3 provides a literature review of the PPS and the frontier assessment methods plus the limited literature on ratio correction.

- Chapter 4 outlines the methodology of this work and visualizes the concept with small examples.

- Chapter 5 studies a real banking example of 811 bank branches based on the new model and reports the findings.
• Chapter 6 provides an analysis and discussion of the bank example results.

• Chapter 7 explains the nonlinearity in more detail and provides a solution.

• Chapter 8 summarizes the key findings, limitation and the prospect of this work.
Chapter 2

RA and DEA

2.1 RA

The efficiency ratio is the traditional measure for productivity in many industries, including banking. At its simplest, it is the cost required to generate each dollar of revenue. Its simplicity is an advantage, but the ratio always needs a business context, and consideration of the complicating factors. Ratio analysis has been in use for a long time and business managers have not lost interest in the method in spite of its proven limitations and consequent errors. Focusing on one factor at a time as the means of comparison is the first and most obvious limitation of RA. Various researchers have shown the shortcomings of RA theoretically as well as pragmatically [Than 96], [Lyon 95]. Krivonozhko et. al [Kriv 08] has recently proved on a rigorous mathematical basis, that ratio analysis misrepresent the performance of the production units under examination.

2.1.1 How RA is measured?

In Ratio Analysis one aspect of performance is selected and then the decision making unit (DMU) having the highest output to input value is labeled with score of one (best performer). The other DMUs which performed below that level are considered inefficient
and their score is measured by dividing their ratio value by the best performer’s ratio value. Consider the output-input ratio metric for input $m$ and output $s$; we can denote the best value based on this metric, corresponding to DMU$_b$, as

$$\frac{y_{sb}}{x_{mb}} = \max_k \left( \frac{y_{sk}}{x_{mk}} \right)$$  \hspace{1cm} (2.1)$$

The RA score for any DMU$_k$ in PPS with input $m$ and output $s$ is, thus, its ratio compared to DMU$_b$:

$$\text{RA}_{(m,s,k)} = \frac{\frac{y_{sk}}{x_{mk}}}{\frac{y_{sb}}{x_{mb}}}$$ \hspace{1cm} (2.2)$$

### 2.1.2 RA and linear programming representation

It was stated by Chen and McGinnis [Chen 07] that RA can be written in a linear program form similar to DEA, here we restate and prove this claim in Theorem 1. However, in RA we are focusing on one input/output at a time.

**Theorem 1** RA$_{(m,s,k)}$: ratio analysis efficiency score of unit $k$, when considering input $m$ and output $s$ and $\theta$ are equivalent,

**Proof:**

$$\theta = \min_{\lambda,\theta} \theta$$ \hspace{1cm} (2.3a)$$

s.t.

$$\sum_{j=1}^{n} \lambda_j \cdot x_{mj} \leq \theta \cdot x_{mk}$$ \hspace{1cm} (2.3b)$$

$$\sum_{j=1}^{n} \lambda_j \cdot y_{sj} \geq y_{sk}$$ \hspace{1cm} (2.3c)$$

$$\lambda_j \geq 0$$ \hspace{1cm} (2.3d)$$

We can divide (2.3b) by $x_{mb}$ and (2.3c) by $y_{sb}$ since both are positive.

$$\theta \cdot \frac{x_{mk}}{x_{mb}} \geq \sum_{j=1}^{n} \lambda_j \cdot \frac{x_{mj}}{x_{mb}}$$

$$\sum_{j=1}^{n} \lambda_j \cdot \frac{y_{sj}}{y_{sb}} \geq \frac{y_{sk}}{y_{sb}}$$
We then can multiply both sides of the above inequality to get:

\[ \theta \cdot \sum_{j=1}^{n} \lambda_j \cdot \frac{y_{sj} \cdot x_{mk}}{y_{sb} \cdot x_{mb}} \geq \sum_{j=1}^{n} \lambda_j \cdot \frac{x_{mj} \cdot y_{sk}}{x_{mb} \cdot y_{sb}} \]

Bringing all to one side and factoring out \( \lambda_j \) we will end up with the following:

\[ \theta \cdot \sum_{j=1}^{n} \lambda_j \cdot \frac{y_{sj}}{y_{sb}} \geq \sum_{j=1}^{n} \lambda_j \cdot \frac{x_{mj}}{x_{mb}} \cdot RA_{(m,s,k)} \]

\[ \sum_{j=1}^{n} \lambda_j (\theta \cdot \frac{y_{sj}}{y_{sb}} - RA_{(m,s,k)} \cdot \frac{x_{mj}}{x_{mb}}) \geq 0 \]

\[ \sum \lambda_j \frac{\theta \cdot y_{sj} \cdot x_{mb} - RA_{(m,s,k)} \cdot x_{mj} \cdot y_{sb}}{y_{sb} \cdot x_{mb}} \geq 0 \] (2.4)

Equation (2.1) implies

\[ \frac{y_{sj}}{x_{mj}} - \frac{y_{sb}}{x_{mb}} = \frac{y_{sj} \cdot x_{mb} - x_{mj} \cdot y_{sb}}{x_{mj} \cdot x_{mb}} \leq 0 \] (2.5)

For Equation (2.5) to hold, Equation (2.4) should be accompanied by:

\[ y_{sj} \cdot x_{mb} \leq x_{mj} \cdot y_{sb} \Rightarrow \theta \geq RA_{(m,s,k)} \] (2.6)

Because if

\[ \theta \leq RA_{(m,s,k)}, \]

then the inequality (2.4) cannot be true. On the other hand, if we find a feasible but not optimal solution to LP (2.3) it will be greater than \( \theta \). One feasible solution to LP (2.3) is choosing \( \lambda = [0, ..., \lambda_b = \frac{y_{sb}}{y_{sb}}, ..., 0] \) and \( \theta = RA_{(m,s,k)} \) and it satisfies all the constraints of the LP (2.3).

\[ \lambda_b \cdot x_{mb} \leq RA_{(m,s,k)} \cdot x_{mk} \]

\[ \lambda_b \cdot y_{sb} \geq y_{sk} \]

As a result

\[ \theta \leq \theta \leq RA_{(m,s,k)} \] (2.7)

From inequalities (2.7) and (2.6) it is concluded that:

\[ \theta = RA_{(m,s,k)} \] (2.8)
2.2 Data Envelopment Analysis

2.2.1 Introduction

The initial DEA model, as originally presented in Charnes, Cooper, and Rhodes (CCR) [Char 78], built on the earlier work of Farrell [Farr 57]. Farrell proposed an activity analysis approach to correct what he believed were deficiencies in the commonly used index number approaches to productivity (and like) measurements. His main concern was to generate an overall measure of efficiency which reflects the measurements of multiple inputs and outputs. Charnes, Cooper, and Rhodes (1978) described DEA as a “mathematical programming model applied to observational data that provides a new way of obtaining empirical estimates of relationships - such as the production functions and/or efficient production possibility surfaces that are the cornerstones of modern economics” [Coop 04]. DEA is a “data oriented” approach for evaluating the performance of a set of peer entities called Decision Making Units (DMUs), which provides a single efficiency score while simultaneously considering multiple inputs and multiple outputs. Because it requires very few assumptions, DEA has opened up possibilities for use in cases which have been resistant to other approaches because of the complex (often unknown) nature of the relationships between the multiple inputs and multiple outputs involved in the operation of the DMUs.

Formally, DEA is a methodology directed to frontiers rather than central tendencies. Instead of trying to fit a regression plane through the data as in statistical regression, for example, one ‘floats’ a piecewise linear surface to rest on top of the observations. Because of this perspective, DEA proves to be particularly adept at uncovering relationships that remain hidden from other methodologies [Coop 04].

Researchers in a number of fields have recognized that DEA is an excellent methodology for modeling operational processes, and its empirical orientation and minimization of a priori assumptions has resulted in its use in a number of studies involving efficient
Chapter 2. RA and DEA

frontier estimation in the nonprofit sector, in the regulated sector, and in the private sector. DEA encompasses a variety of applications in evaluating the performances of different kinds of entities, such as hospitals, universities, cities, courts, business firms, and banks among others. According to the recent bibliography [Emro 08], over 4000 papers were published on DEA by 2007. Such rapid growth and widespread (and almost immediate) acceptance of the methodology of DEA is a testimony to its strengths and applicability.

2.2.2 DEA Basic assumptions and Models

We have talked about the history of DEA, the motivation behind it and the freedom it offers compared to statistical methods. Here we look at the principles and assumptions around various DEA models that we use in this work.

2.2.3 Production Possibility Set

In productivity analysis, or efficiency measurement in general, when the DMUs consume $s$ different inputs to produce $m$ different outputs, the production possibility set is the collection of all feasible DMUs that are capable of producing output $y = (y_1, y_2, \ldots, y_m)$ consuming input $x = (x_1, x_2, \ldots, x_s)$. The PPS is defined as the set:

$$
\Psi = \{(x, y) \in R^{m+s} \| x \text{ can produce } y\}
$$

As mentioned in section 2.2.1, DEA is very data oriented. It means that we build the production possibility set based on observed data points and some assumptions which, in some aspects, relates to our model. We briefly introduce some of the assumptions used but leave the detailed evaluation and the choice of the appropriate model for the next chapter.

**Disposability axiom:** A fundamental assumption to form the PPS out of the available data is ‘disposability’. If $x$ can produce $y$ so does any $x_i \geq x$ and if $y$ could be produced by
Chapter 2. RA and DEA

$x$ so could be any $y_j \leq y$. Formally each observed data $X = (x_1, ..., x_m), Y = (y_1, ..., y_s)$ brings along part of the unobserved piece of the PPS which is defined as

$$\{(X^i, Y^i) \in R^{m+s}| X^i \geq X \text{ and } Y^i \leq Y\}$$

This is like saying if DMU$_i$ could be realized then any DMU that is doing worse is feasible, too. This assumption leads to the Free Disposal Hull (FDH) model, which shares its PPS with most of the other models.

**Convexity:** Any convex linear combination of realized DMUs is feasible. In other words if two DMUs are in the PPS so does the line connecting these two ones (any linear combination of them). DMU$_{\text{composite}} = \{\sum_{i=1}^{n} \lambda_i \cdot \text{DMU}_i | \sum_{i=1}^{n} \lambda_i = 1\}$. This assumption leads to the BCC model, a variable return to scale model.

**Ray Unboundedness:** Scaling up or down of any realized DMU generates a new feasible DMU. $\{\forall \text{DMU}_i \in \text{PPS and } \gamma \geq 0, \gamma \cdot \text{DMU}_i \in \text{PPS}\}$. This assumption added to the convexity assumption is the basis of CCR, a constant returns to scale model.

### 2.2.4 Frontier

Once we generate the desired PPS, set $\psi$ in 2.9, then it is time to define the potential benchmarks or the frontier. The frontier is composed of one or more estimated lines or surfaces (depending on dimensions) enveloping only but no less than the whole PPS. It is the line between feasibility and in-feasibility of the DMUs. We might be interested in certain facets of the frontier, depending on the models’ orientation. The projection to the frontier maybe to an input or output facet (segment) of the frontier.

$$\partial \Psi_x = \{y|(x, y) \in \Psi, (x, \eta \cdot y) \notin \Psi, \forall \eta > 1\}$$

$$\partial \Psi_y = \{x|(x, y) \in \Psi, (\theta \cdot x, y) \notin \Psi, 0 < \theta < 1\}$$

Figure 2.1 shows different frontiers based on disposability, convexity and ray unboundedness assumptions on PPS.
2.2.5 Efficiency Definition

What do we mean by “efficiency”, or more generally, by saying that one DMU is more efficient than another DMU? Relative efficiency in DEA provides us with the following definition, which has the advantage of avoiding the need for assigning a-priori measures of relative importance to any input or output.

**Full Efficiency:** Full efficiency is attained by any DMU if and only if none of its inputs or outputs can be improved without worsening some of its other inputs or outputs. In most management or social science applications the theoretically possible levels of efficiency will not be known. The preceding definition is therefore replaced by emphasizing its uses with only the information that is empirically available as in the following definition.

**Full Relative or Technical efficiency:** Full technical efficiency is attained by any DMU if and only if, compared to other observed DMUs, none of its inputs or outputs can be improved without worsening some of its other inputs or outputs. Relative efficiency of
the unit when compared to its peer groups is the amount of input that could be eliminated or the amount of output increased without worsening any other input or output. The peer groups might change over time or as the result of managerial decisions, because of changing the production technology or merging with other entities.

**Technical change:** Technical change is the relative efficiency of the entity when compared to a broader or newer peer groups. It represents the difference of the organization’s environment and technology adoption or, technically speaking, the benchmark shift [Grif 99]. Benchmarks can be different as the result of our choice of models and assumptions. We might assume our business is scalable with no overhead cost or assume a minimum amount of input to start production.

**Scale efficiency:** Banker et al. [Bank 84] identify the difference between the “variable returns to scale” model, BCC, and the “constant return to scale” model, CCR, as a production scale effect. Scale efficiency represents the failure in achieving the most productive scale size and the score difference between CCR and BCC models reflects that. It is computed as the CCR efficiency score divided by the BCC efficiency score: \[ \frac{\theta_{\text{CCR}}}{\theta_{\text{BCC}}} \]. In figure 2.1 for DMU “D” the difference between “\( t_2 \)” and “\( t_3 \)” is because of the scale inefficiency.

**Input Slack factor:** identifies the minimum value \( x \) for input \( m \) without changing other inputs or output \( y \) when \( (x, y) \) belongs to the PPS

**Input substitution factor:** identifies the smallest value \( x \) for input \( m \) that is possible for any \( x \) such that \( (x, y) \) belongs to the PPS

**Output Slack factor:** identifies the maximum value \( y \) for output \( s \) without changing other outputs or inputs when \( (x, y) \) belongs to the PPS

**Output substitution factor:** identifies the largest value \( y \) for output \( s \) that is possible for any \( y \) such that \( (x, y) \) belongs to the PPS
2.2.6 Orientation

DMUs are represented by their inputs and outputs. Efficiency scores depend on how far the DMU is located from the frontier. Depending on the problem, DMUs can reduce their inputs or increase their outputs or target improvement in inputs and outputs simultaneously in order to move to a point on the frontier.

2.2.7 Additive model

While CCR and BCC models are either focused on minimizing input (input oriented) or maximizing output (output oriented), Additive model is the one that focus on decreasing input and increasing output simultaneously and therefore has no orientation. Additive model shares the same PPS with BCC model. In figure 2.1 point “E” is an inefficient DMU and in input orientation BCC model, “E” needs to reduce inputs and some output slacks to reach to “A”, same analogy is true for output orientation which leads “E” to “G”. However in Additive model point “B” is optimum because reaching to that point requires overall maximum cuts in waste and shortfalls.

2.3 Linkage between Data Envelopment Analysis and Ratio Analysis

A number of researchers have studied DEA and RA and noted the positive and negative differences between them. While some papers compared DEA and RA, others attempted to combine or relate the two methods.

2.3.1 Comparing DEA and RA

Cronje [Cron 02] compared the use of the DuPont system with DEA in measuring profitability of local and foreign controlled banks in South Africa. The DuPont system
[Gall 03] is an analysis technique to determine what processes the company does well and what processes can be improved by focusing on the interrelationship between return on assets, profit margins and asset turnover. The results show that DEA gives a more accurate classification because it provides a combined comparison of the performance of the banks with regard to different financial ratios beyond the three ratios involved in the DuPont system.

Feroz et al. [Fero 03] tested the null hypothesis that there is no relationship between DEA and traditional accounting ratios as measures of performance of a firm. Their results reject the null hypothesis indicating that DEA can provide information to analysts that is additional to that provided by traditional ratio analysis. They applied DEA to the oil and gas industry to demonstrate how financial analysis can employ DEA as a complement to ratio analysis.

Thanassoulis et al. [Than 96] studied DEA and performance indicators (output to input ratios) as alternative instruments of performance assessment using data from the provision of perinatal care in England. They compared the two aspects of performance measurement and target settings. As performance measures are concerned, in a typical multi-input multi-output situation, various ratios should be defined. But, this makes it difficult to gain an overview of the unit’s performance particularly when the different ratios of that unit do not agree on the unit’s performance, as is often the case. Yet selecting only some ratios can bias the assessment. They found their DEA and individual ratios agree weakly on unit performance and this is because DEA reflects the overall efficiency while RA only the specific ones. On the second aspect, namely target setting, DEA identifies input/output levels which would render a unit efficient. Ratio based targets may result in unrealistic projections because they are derived with reference to one input and one output at a certain point in time without regard for the rest of the input-output levels. But the authors believe that ratios could give some useful guidance for further improvements of the efficient units in DEA.
Bowlin [Bowl 04] used DEA and ratio analysis to assess the financial health of companies participating in the Civil Reserve Air Fleet – an important component of the Department of Defense’s airlift capability – over a 10 year period. He employed DEA and then tried to explain the observations based on ratio analysis. He believes the two methods together gave insight to the study.

2.3.2 Combining DEA and RA

There are other studies which attempted to combine or relate the two methods. As we proved in section 2.1.1, ratio analysis (RA) is the same as the CCR model when the DMUs have only a single input and a single output.

For example Chen and Agha [Chen 02b] characterized the inherent relationships between the DEA frontier DMUs and output-input ratios. They showed that top-ranked performance by ratio is a DEA frontier point and DEA subsumes the premise of the RA, however, it fails to identify all types of dominating units as DEA does.

Gonzalez-Bravo [Gonz 07] proposed a Prior-Ratio-Analysis procedure which is based on the existence of a relationship of individual ratios to DEA efficiency scores. He listed the efficient units whose efficiencies are overestimated by DEA because they perform highly in a single dimension and the inefficient units whose efficiency are underestimated because they perform reasonably well in all the considered dimensions but do not stand out in any of them.

Being motivated by Chen and Agha’s [Chen 02b] paper, Wu and his colleagues [Wu 05] proposed an aggregated ratio analysis model in DEA. This ratio model has been proven to be equivalent to the CCR model. However, we believe the inclusion of all possible ratios in the model does not necessarily make sense and the number of possible ratios grows in size exponentially as the number of inputs-outputs increase. To illustrate this, consider a three input three output case; we will have \((2^3 - 1)\) aggregated inputs and \((2^3 - 1)\) aggregated outputs, so the model optimizes \(7 \times 7 = 49\) aggregated ratios – where
some of them do not represent a meaningful concept such as the ratio of cured people in ICU to the number of kids admitted in ER. The authors have also proven that a subset of all the possible aggregated ratios is also equivalent to the CCR model and for our example the number of variables decrease significantly, which is a substantial improvement. However, the necessities of including unrelated ratios still remain unaddressed. To deal with some meaningless ratios Despić et al. [Desp 07] proposed a DEA-R efficiency model, in which all possible ratios (output/input) are considered as outputs. This model enables the analyst to easily translate some of the expert opinions into weight restrictions in terms of ratios. This creates an immediate communication between the experts and the model.

In another interesting study Chen and McGinnis [Chen 07] showed there is a bridge between ratio efficiency and technical efficiency. Therefore, $RA_{(m,s,k)}$ is a product of seven different component measurements: technical efficiency, technical change, scale efficiency, input slack factor, input substitution factor, output slack factor and output substitution factor. Technical efficiency, technical change and scale efficiency are DMU dependent only, i.e., for a DMU$_k$, they will be the same no matter what input $m$ and output $s$ are selected for $RA_{(m,s,k)}$. Input slack factors and input substitution factors are DMU and input dependent. For a particular $RA_{(m,s,k)}$, they depend on the selection of input $m$ and DMU$_k$ but not output $s$. However, output slack and substitution factors are functions of DMU$_k$, input $m$ and output $s$. This relationship provides a basis for concluding that the conventional partial productivity metric is not a proper performance index for system benchmarking. This is because it depends upon other effects in addition to the system-based technical efficiency between a given DMU and a “benchmark” DMU. Furthermore, $RA_{(m,s,k)}$ is the product of Technical efficiency and the other six factors which are all less than or equal to one. Therefore, $RA_{(m,s,k)}$ being close to one indicates that all seven factors should be close to one, and, in fact, larger than $RA_{(m,s,k)}$. This property partly explains why the DMUs with the largest output-input ratio will be technically efficient.
when their RA equals to one [Chen 02a].

Other researchers started to use the financial ratios as inputs and outputs in DEA, with an expectation that they can get the best out of that. In a Magyar Nemzeti Bank Working Paper, Holló and Nagy [Holl 06] employed ratios in their production model to assess 2459 banks in the European Union. Hollingsworth and Smith, for the first time, pointed out the inaccuracy of the CCR model [Holl 03], when data are in the form of ratios. Then Emrouznejad et al. [Emro 09] examined the problem of ratios in more detail and proposed a series of modified DEA models and our research is built upon one of their models. We will later explain, in full detail, why the other models were ruled out. Our proposed second phase to maximize the slacks and how we tackle the nonlinear model will also be explained.

2.4 Advantages of incorporating ratios in DEA

But why would incorporating ratios in DEA be beneficial? Some of the benefits of using ratios are: they will reduce the number of variables, which becomes important when we do not have a large pool of DMU data; they also enable us to consider highly correlated variables as a one meaningful ratio; but the most important reason is that it makes the DEA analysis reporting more comprehensible to senior management who are used to ratios and they might better see the benefits of DEA in efficiency evaluations.
Chapter 3

Frontier Assessment

In the previous chapter PPS and the efficiency frontier were explained briefly. Since the methodology of this research is mainly based on the notion of a PPS and frontier assessment methods, this chapter is a literature review of the different approaches to approximate the efficiency “frontier”.

All DMUs would belong to the subspace between the origin and the frontier or the frontier and infinity depending on output or input orientation respectively. The concept of a frontier is more general and easier to understand than the concept of a “production function” which has been regarded as a fundamental concept in economics, because the frontier concept admits the possibility of multiple production functions, one for each DMU, with the frontier boundaries consisting of hypothetical DMUs which are created out of the more efficient members of the set. Each DMU is projected to the part of the benchmark frontier that has the same or similar production pattern, so it is vital to get that benchmark right because otherwise we would set a false goal for the DMU under study. Hence, we might either underestimate or overestimate the efficiency score and, as a result, give an unrealistic projection.
3.1 Deterministic Frontier Estimators

3.1.1 Free Disposal Hull Frontier

The Free Disposal Hull assumption adds the unobserved production points with output levels equal to or lower than those of some observed points and more of at least one input; or with input levels equal to or higher than those of some observed points and less of at least one output to the observed production data [Depr 84]. In other words if $x$ generates $y$ then more $x$ can still generate $y$ and $x$ can generate less $y$, too. FDH is assumed, in the literature, to be sufficient to induce a reference set that has all the properties economic theory requires of a production set [Tulk 93]. However, strong disposability assumptions exclude congestion, which is frequently observed e.g. in agriculture and transportation and undesired outputs (or inputs) e.g. in oil production. As a simple example, if 100 trucks can deliver goods in a specific route, within a certain time, then 1000 trucks (more input) might not necessarily perform at the same level because the entire route might not have the capacity to handle 1000 trucks. Assessment of congestion analysis within DEA and ways to deal with it could be found in a work by Färe et al. [Fare 83], Brockett et al.[Broc 98] and Cherchye et al. [Cher 01]. For more information on the undesired output and disposability issue, refer to Hongliang and Pollitt’s work on environmental efficiency [Yang 07]. The FDH frontier looks like a staircase for the one input one output case as seen in figure 3.1.

3.1.2 Various Returns to Scale Frontiers

The convexity assumption adds any non-observed data that is a convex combination of some points in the FDH to the PPS. Although there are notable arguments and evidence favouring convexity, some researchers have found this axiom very restrictive and proposed to drop or weaken it. For a complete study on this issue see Cherchye et al. [Cher 99].
3.1.3 Constant Returns to Scale Frontier

The full proportionality assumption includes any non-observed production point that is proportional to some data points in the FDH. This assumption is in accordance with the original DEA model (CCR) and the Farrel efficiency measure. It is critical to know that in DEA we assume that a linear combination of DMUs is possible and real. The CCR frontier contains the other frontiers and so if a DMU be on CCR frontier it will be on the FDH and BCC frontiers also, obviously the reverse is not true.

3.2 Probabilistic Frontier Estimators

The nonparametric deterministic estimators envelop all the data points and so are very sensitive to noise. They may be seriously affected by the presence of outliers (super-efficient observations), as well as data errors, which may lead to a substantial underestimation of the overall efficiency scores. Therefore, in order to assure credibility of the efficiency indices, it is crucial to adopt some additional method to correct for such discrepancies; only then may one hope to obtain estimators that could be useful for the decision making process. In probabilistic models we look at the PPS with some doubt
and leave some room for errors and randomness. In the production process: $x$ produces $y$ is defined with a joint probability. For $(x,y)$ to be in the PPS it is essential to have:

$$F(x,y) = \{\text{Prob}(X \geq x, Y \leq y)\} > 0 \text{ be a dominant point}$$

$$H(x,y) = \{\text{Prob}(X \leq x, Y \geq y)\} > 0 \text{ be a dominated point}$$

Based on the $n$ observed data points, the FDH estimator of the PPS, the frontier and the efficiency score estimator for an input-oriented case are:

$$\hat{F}_n(x,y) = \sum_{i=1}^{n} \text{Prob}(X_i \geq x, Y_i \leq y) \quad (3.1)$$

$$\hat{H}_n(x,y) = \sum_{i=1}^{n} \text{Prob}(X_i \leq x, Y_i \geq y) \quad (3.2)$$

$$\hat{\theta}_n(x,y) = \inf\{\theta | \hat{F}_n(\theta \cdot x|y) > 0\} \quad (3.3)$$

$$\hat{\theta}_n(x,y) = \inf\{\theta | \hat{H}_n(\theta \cdot x|y) > 0\} \quad (3.4)$$

It has been proven by Park et al. [Park 00] that $\hat{\theta}(x,y)$ is a consistent estimator of $\theta(x,y)$ with the convergence rate of $n^{-1/(m+s)}$ where $m=$number of inputs and $s=$number of outputs.

### 3.2.1 Partial m Frontier

Cazals et al. [Caza 02] introduced the concept of partial frontiers (order-m frontiers) with a nonparametric estimator which does not envelop all the data points. While keeping its nonparametric nature, the expected order-m frontier does not impose convexity on the production set and allows for noise (with zero expected values). For example, to measure the input efficiency of $(x,y)$, we pick $m$ random DMUs with inputs $X_i$ and output at the same level or better than $y$, then we estimate the FDH PPS and efficiency score based on those $m$ DMUs. So $(x,y)$ is compared to a set of $m$ peers producing more than its level $y$ and we take this as the benchmark, the expectation of the minimal achievable input in
place of the absolute minimal achievable input.

\[ \tilde{\Psi}_m(y) = \{(x^*, y^*)|x^* \geq X_i, y^* \leq y\} \]

\[ \tilde{\theta}_m(x, y) = \inf \{\theta|(\theta \cdot x, y) \in \tilde{\Psi}_m(y)\} \]

Since \( m \) DMUs are random variables, so are \( \tilde{\Psi}_m(y) \) and \( \tilde{\theta}_m(x, y) \), so the input efficiency score, on average, is:

\[ \hat{\theta}_m(x, y) = \mathbb{E}(\tilde{\theta}_m(x, y)|Y \geq y^*) \quad (3.5) \]

Therefore, instead of looking for the lower boundary, input orientation frontier, the order-\( m \) efficiency score can be viewed as the expectation of the minimum input efficiency score of the unit \((x, y)\), when compared to \( m \) units randomly drawn from the population of units producing more outputs than the level \( y \). This is certainly a less extreme benchmark for the unit \((x, y)\) than the absolute minimal achievable level of inputs. The order-\( m \) efficiency score is not bounded by one: a value greater than one indicates that the unit operating at the level \((x, y)\) is more efficient than the average of \( m \) peers randomly drawn from the population of units \((n \text{ observed DMUs})\) producing more output than \( y \).

\[ \hat{\theta}_{m,n}(x, y) = \hat{\theta}_n(x, y) + \int_{\hat{\theta}_n(x, y)}^{\infty} (1 - F_X(u|x|y)^m)du \]

\[ \lim \hat{\theta}_{m,n}(x, y) = \hat{\theta}_n(x, y), \text{as} \ m \to \infty \]

For a finite \( m \) the frontier may not envelope all data points. The value of \( m \) may be considered as a trimming parameter and as \( m \) increases the partial order-\( m \) frontier converges to the full-frontier. It is shown that by selecting the value of \( m \) as an appropriate function of \( n \), the non-parametric estimator of the order-\( m \) efficiency scores provides a robust estimator of the corresponding efficiency scores sharing the same asymptotic properties as the FDH estimators, but being less sensitive to outliers and/or extreme values. In the literature numerical methods like the Monte-Carlo procedure is being used instead of evaluating multivariate integrals. In Chapter 7 of this work we will come back to the idea of m-frontiers and build our own Monte-Carlo method to derive the frontier estimator when we face nonlinearity.
3.2.2 Quantile Frontier

Aragon et al. [Arag 05] proposed an alternative approach to order-m partial frontiers by introducing quantile based partial frontiers. The idea is to replace the concept of the “discrete” order-m partial frontier by a “continuous” order $\alpha$ partial frontier where $\alpha \in [0, 1]$ corresponds to the level of an appropriate non-standard conditional quantile frontier. This method is more robust to the effects of outliers. The original $\alpha$-quantile approach was limited to one dimensional input for the input oriented frontier and to one dimensional output for the output oriented frontier, however, Daouia and Simar [Daou 07] developed the $\alpha$-quantile model for multiple inputs and outputs. From equation 3.4

$$\tilde{\theta}(x, y) = \inf\{\theta | H(\theta \cdot x, y) > 0\}$$

$\alpha$ - quantile input efficiency is defined as

$$\tilde{\theta}_\alpha(x, y) = \inf\{\theta | H(\theta \cdot x, y) > 1 - \alpha\}$$

(3.6)

Unit $(x, y)$ consumes less than $\alpha100\%$ of all other units producing output larger than or equal to $y$ and consumes more than the $(1 - \alpha)100\%$ remaining units. If $\theta(x, y) = 1$, we will say that the unit is input efficient at the level $\alpha100\%$. Clearly when $\alpha = 1$, this is the same as the Farrell-Debreu input efficiency score sharing the same properties of the FDH estimator, but since it does not envelop all the data points, it will be more robust to extreme and/or outlying observations [Daou 07].

3.2.3 Practical Frontiers

As we have seen, DEA is very data oriented and it builds the PPS based on certain assumptions. DEA does not have any benchmark to rank the efficient units against and, since its vision is limited to the sampled data, it cannot perceive any potential improvement beyond the already identified efficient DMUs. Moreover, although the fundamental assumptions hold, on average, in practical cases there might be exceptions for some entities due to either managerial or natural restrictions. For example, although we believe
that any linear combination of DMUs could be realized, we do not know if any inefficient DMU projected to a target, can imitate that production by changing its inputs/outputs accordingly. One of DEA’s limitations is associated with its inability to provide any further insight into the DMUs on the frontier. However, there might be a possibility for the DEA efficient DMUs to improve and it is important for management to set targets for their efficient units if the organization is to advance as a whole. Sowlati and Paradi [Sowl 04] looked at the problem and formed a new practical frontier, by possible changes in the inputs/outputs of the already efficient DMUs. They introduced a novel linear programming approach to create those hypothetical DMUs, and formed a new practical frontier. Other researchers have worked on a practical frontier by introducing weight restrictions on inputs/outputs to prevent DEA from setting a practically impossible target on the frontier for an inefficient unit. The 73rd Annals of Operations Research was dedicated to “extending the frontiers of DEA” and it includes various papers on the issue [Lewi 97]. In our work, we will use numerical methods to generate hypothetical DMUs to build the practical frontier which dominates the conventional DEA frontier.
Chapter 4

Incorporating ratios in DEA

When inputs and/or outputs are in the form of ratios, conventional linear programming methods may fail to build the correct frontier and, as a result, the scores and projections are distorted. We reviewed the limited literature on the “ratios” problem in Section 2.3.2 and promised to investigate the two existing modified models by Emrouznejad et al.[Emro 09], and explain why our research is built upon one of their models.

4.1 Modified DEA models

Recall that the main concern about using ratios as input or output variables in the context of the conventional DEA is due to the fact that DEA estimates the PPS out of the $n$ available data points. Conventional DEA identifies each DMU by its production process $(x_i, y_i)$ and works with a linear combination of $(x_i, y_i)$s. As long as $x_i$ and $y_i$ do not contain any ratios, working with inputs and outputs translates exactly into working with DMUs. However, merging inputs and outputs is not equivalent to merging DMUs when ratios are involved. For example the output of a composite DMU, $Y^*_r$ in 4.2 and the composite output, $Y_r$ in 4.1 are not equal.
Hollingsworth et al. [Holl 03] believed that the BCC formulation is the appropriate model when ratios are involved because: \( Y_r^* = Y_r \Rightarrow \sum \lambda_i = 1 \). However, we found the proof presented in the paper does not justify their claim, because \( \sum \lambda_i = 1 \) does not guarantee \( Y_r^* = Y_r \). In other words, it is more precise to state: for the DEA approach to be valid when using ratios, the BCC model is a necessary condition, but not a sufficient one. Emrouznejad [Emro 09] examined the problem and suggested two models: The first breaks down ratios to their numerator and denominator parts and treat one part as input and the other as an output. The second one is based on the idea of working with DMUs rather than their production processes.

### 4.1.1 Model I

In this model each output in the ratio form has been split up into two parts: denominator and numerator and each part has been dealt with separately. For example when the output, in ratio form, is meant to be maximized the numerator is maximized while the denominator is minimized. This is done through adding one input and one output for
any ratios we have.

\[
\begin{align*}
    \min_{\lambda, \theta} & \quad \theta \\
    \text{s.t.} & \quad \sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq \theta \cdot x_{ik}, \quad i = 1, \ldots, m \tag{4.3a} \\
    & \quad \sum_{j=1}^{n} \lambda_j \cdot y_{ij} \geq y_{ik}, \quad i = 1, \ldots, s, i \neq r \tag{4.3b} \\
    & \quad \sum_{j=1}^{n} \lambda_j \cdot n y_{rj} \geq n y_{rk}, \quad i = r \tag{4.3c} \\
    & \quad \sum_{j=1}^{n} \lambda_j \cdot d y_{rj} \leq \theta \cdot d y_{rk} \tag{4.3d} \\
    & \quad \lambda_j \geq 0 \tag{4.3f}
\end{align*}
\]

The obvious disadvantage of this model is that it increases the number of variables. This might create difficulties when the number of outputs and inputs becomes large compared to the number of DMUs. More precisely the general rule for DEA models requires \( n \geq \max\{3(m + s), m \times s\} \). Because any output in the form of a ratio adds one additional input, the \( n \geq \max\{3(m + r + s), (m + r) \times s\} \) limits the number of ratio variables to \( r \leq \min\{n/3 - (m + s), n/s - m\} \). Besides that, we argue that this model does not capture the disposability of output in the ratio form because while \( n y_{ru}/d y_{ru} \geq n y_{rf}/d y_{rf} \) may hold, \( n y_{ru} \geq n y_{rf} \) or \( d y_{ru} \leq d y_{rf} \) may not. i.e. \( 1/2 \leq 3/5 \) but the DMU with output of \( 1/2 \) will not belong to the PPS because \( 2 < 5 \).

The other issue is correlation. Usually when we introduce ratios the numerator and denominator are highly correlated, e.g. the revenue from commercial loans and the amount of money dedicated to the commercial loans. The definition of input orientation is in jeopardy too, because it is a mixed orientation here.

Sometimes the nature of the denominator is discretionary. While we cannot change the discretionary (e.g. environmental) variables, they reflect some characteristics of the DMU and in the form of a ratio their impact is acknowledged. With model I, imposing
a constraint on discretionary variable is problematic. Discussion about adding environmental variables is beyond the scope of this work.

4.1.2 Model II

The other model proposed by Emrouznejad follows from the correct understanding of convexity. Usually, in frontier analysis and DEA, the PPS is said to be convex. This is typically shown by the input output sets such as:

\[
\tilde{\Psi}_{DEA} = \{ (x, y) \in R^{m+s} | y \leq \sum_{i=1}^{n} \lambda_i y_i, x \geq \sum_{i=1}^{n} \lambda_i x_i & \sum_{i} \lambda_i = 1 \} 
\]

\[
\hat{\Psi}_{DEA} = \{ U \in R^{m+s} | y_U \leq y \sum_{i=1}^{n} \lambda_i U_i, x_U \geq x \sum_{i=1}^{n} \lambda_i x_i & \sum_{i} \lambda_i = 1 \}
\]

If the inputs and outputs do not include any ratios the two sets are equal. Otherwise we need to stick to the general convexity axiom as was stated in the second set. Without any loss of generality, we assume the first \(r\) outputs are in the form of ratios. Then the LP for the BCC input orientated case will be:

\[
\text{min } \theta \\
\lambda, \theta \\
\text{s.t.} \\
\sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq \theta \cdot x_{ik} \quad i=1...m \\
\sum_{j=1}^{n} \lambda_j \cdot n y_{ij} \geq y_{ik} \sum_{j=1}^{n} \lambda_j \cdot d y_{ij} \quad y_{ik} = n y_{ik} / d y_{ik}, \quad i=1...r \\
\sum_{j=1}^{n} \lambda_j \cdot y_{ij} \geq y_{ik} \quad i = r + 1 \ldots s \\
\sum_{j=1}^{n} \lambda_j = 1 \\
\lambda_j \geq 0
\]
which is the envelopment form of

\[
\theta(DMU_k) = \max \sum_{i=r+1}^{s} u_i y_{ik} - u_0 \quad (4.5a)
\]

\[
\sum_{i=1}^{m} v_i x_{ik} = 1 \quad (4.5b)
\]

\[
\sum_{i=1}^{r} w_i (n y_{ij} - y_{ik} d_{ij}) + \sum_{i=r+1}^{s} u_i y_{ij} - \sum_{i=1}^{m} v_i X_{ij} - u_0 \leq 0 \quad \forall j \quad (4.5c)
\]

\[
w_i, u_i, v_i \geq 0, u_0 \text{ free} \quad (4.5d)
\]

Here we see that the price for the ratios do not appear in the objective function but they play an indirect role in order to maximize the efficiency score. If the DMU is good in any ratio, the first term in the second constraint becomes negative which implies that the constraint is less binding and the score can grow. But if the DMU is not a good performer in some ratios then the constraints become more binding and that may result in a lower score, which we expect.

**Seeking the Maximum Slack**

Although model 4.4 was suggested in the literature, there has been no work on implementing the model. In DEA an entity is efficient if and only if the efficiency score is one and the slacks are zero. In their book, Cooper et al. [Coop 04] suggested a two phase DEA that fulfills the search for efficient DMUs without requiring the non-Archimdean element. The optimization LP 4.4 represents phase I. For phase II we have developed
the following:

\[
\begin{align*}
\max_{\lambda_i, \sigma_i} & \sum_{i=1}^{m} s_i^- + \sum_{i=1}^{s} s_i^+ \quad \text{(4.6a)} \\
\text{s.t.} & \\
\sum_{j=1}^{n} \lambda_j \cdot x_{ij} - \theta \cdot x_{ik} + s_i^- = 0 & \quad i = 1..m \quad \text{(4.6b)} \\
\sum_{j=1}^{n} \lambda_j \cdot y_{ij} - y_{ik} - s_i^+ = 0 & \quad \forall i = r + 1...s \quad \text{(4.6c)} \\
\frac{\sum_{j=1}^{n} \lambda_j \cdot ny_{ij}}{\sum_{j=1}^{n} \lambda_j \cdot dy_{ij}} - y_{ik} - s_i^+ = 0 & \quad y_{ik} = ny_{ik}/dy_{ik} \quad \forall i = 1...r \quad \text{(4.6d)} \\
\sum_{j=1}^{n} \lambda_j = 1 & \quad \text{(4.6e)} \\
\lambda_j \geq 0 & \quad \text{(4.6f)}
\end{align*}
\]

Clearly, the constraint 4.6d is nonlinear, we linearize it in two steps by transforming variables and introducing two new variables \(\Delta\) and \(\omega\):

\[
\begin{align*}
1 \cdot \sum_{j=1}^{n} \lambda_j \cdot (ny_{ij} - dy_{ij} \cdot y_{ik}) - \sum_{j=1}^{n} \lambda_j \cdot dy_{ij} \cdot s_i^+ = 0 & \quad \forall i = 1...r \\
\omega_{ij} = ny_{ij} - dy_{ij} \cdot y_{ik} & \quad \lambda_j \cdot s_i^+ = \Delta_{ij} \\
2 \cdot \sum_{j=1}^{n} \lambda_j \cdot w_{ij} - \sum_{j=1}^{n} \Delta_{ij} \cdot dy_{ij} = 0 & \quad \forall i = 1...r
\end{align*}
\]

To simplify the notation we can assume that the non-ratio outputs are ratios with de-
nominator 1 and the linearized LP form will appear as:

\[
\begin{align*}
\text{max} & \quad \lambda, s_i, \Delta_{ij} \left( \sum_{i=1}^{m} s_i^- + \sum_{j=1}^{n} \sum_{i=1}^{s} \Delta_{ij} \right) \\
\text{s.t.} & \quad \sum_{j=1}^{n} \lambda_j \cdot x_{ij} - \theta \cdot x_{ik} + s_i^- = 0 \quad i=1...m \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot w_{ij} - \sum_{j=1}^{n} \Delta_{ij} \cdot dy_{ij} = 0 \quad i = 1...s \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \\
& \quad \lambda_j \geq 0
\end{align*}
\] (4.7a, 4.7b, 4.7c, 4.7d, 4.7e)

Linearized LP 4.7 comes at a cost of increasing the output slack variables “r” times (r is the number of ratio variables). However, the number of variables in contrast to the number of constraints have not been a significant issue with respect to computational complexity. In fact, computation time tends to be roughly proportional to the cube of the functional constraints while the number of variables is a relatively minor factor. For more on the simplex complexity please refer to Section 4.8 of Hillier’s book [Hill 01].

### 4.2 Motivating Example

We summarize what has been said so far about DEA and ratios with a simple example. Assume that the DMUs are branches of a company engaged in selling laser equipment for eyesight correction. The data is in table 4.1. In practice, management prefers to see the information in a more comprehensive way as shown in table 4.2.

Following the DEA approach and evaluating each branch’s performance (here we evaluate \(branch_B\)), we search for a hypothetical DMU which outperforms \(branch_B\). If we
### Chapter 4. Incorporating ratios in DEA

<table>
<thead>
<tr>
<th>Per season</th>
<th>Branch A</th>
<th>Branch B</th>
<th>Branch C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Returning customers</td>
<td>30</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>Technical Staff hours</td>
<td>600</td>
<td>300</td>
<td>200</td>
</tr>
<tr>
<td>Number of units sold</td>
<td>130</td>
<td>144</td>
<td>96</td>
</tr>
<tr>
<td>Sales Staff hours</td>
<td>650</td>
<td>1200</td>
<td>800</td>
</tr>
<tr>
<td>Commercial Expenses</td>
<td>100</td>
<td>70</td>
<td>110</td>
</tr>
</tbody>
</table>

*Table 4.1: Lasik Equipment Information*

<table>
<thead>
<tr>
<th>output indicator</th>
<th>output indicator</th>
<th>input indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Satisfaction:</td>
<td>Revenue generation:</td>
<td>Cost</td>
</tr>
<tr>
<td>Number of returning customers / Technical Staff hours</td>
<td>Number of units sold / Sales Staff hours</td>
<td></td>
</tr>
<tr>
<td>Branch A 20%</td>
<td>5%</td>
<td>100</td>
</tr>
<tr>
<td>Branch B 12%</td>
<td>8%</td>
<td>70</td>
</tr>
<tr>
<td>Branch C 5%</td>
<td>12%</td>
<td>110</td>
</tr>
</tbody>
</table>

*Table 4.2: Comprehensive Lasik Equipment Information*
focus on the production process as in table 4.2, the LP for \(branch_B\) is:

\[
\begin{align*}
\text{max } & \eta_B & \quad (4.8a) \\
\text{s.t.} & \\
\lambda_1 20\% + \lambda_2 12\% + \lambda_3 5\% & \geq \eta_B \times 12\% & \quad (4.8b) \\
\lambda_1 5\% + \lambda_2 8\% + \lambda_3 12\% & \geq \eta_B \times 8\% & \quad (4.8c) \\
\lambda_1 100 + \lambda_2 70 + \lambda_3 110 & \leq 70 & \quad (4.8d) \\
\sum_i \lambda_i & = 1 & \quad (4.8e) \\
\lambda_i & \geq 0 & \quad (4.8f)
\end{align*}
\]

However, if we switch the focus to the branches, \(DMU_{\text{hypothetical}} = \lambda_1 \cdot \text{branch}_A + \lambda_2 \cdot \text{branch}_B + \lambda_3 \cdot \text{branch}_C\). The hypothetical DMU gets some share of each peer DMU,

<table>
<thead>
<tr>
<th>output indicator</th>
<th>output indicator</th>
<th>input indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Satisfaction :</td>
<td>Revenue generation:</td>
<td>Cost</td>
</tr>
<tr>
<td>Number of returning customers</td>
<td>Number of units sold</td>
<td></td>
</tr>
<tr>
<td>Technical Staff hours</td>
<td>Sales Staff hours</td>
<td></td>
</tr>
<tr>
<td>(\frac{\lambda_1 30 + \lambda_2 24 + \lambda_3 96}{\lambda_1 600 + \lambda_2 300 + \lambda_3 800})</td>
<td>(\frac{\lambda_1 130 + \lambda_2 144 + \lambda_3 10}{\lambda_1 650 + \lambda_2 1200 + \lambda_3 200})</td>
<td>(\lambda_1 100 + \lambda_2 70 + \lambda_3 110)</td>
</tr>
</tbody>
</table>

Table 4.3: Hypothetical DMU Information

meaning a percent share of the “number of the people who work in each branch”, the “revenues” and so on. The hypothetical DMU does not get the same share of the ratios,
although the LP for \( branch_B \) is:

\[
\begin{align*}
\text{max} & \quad \eta_B \\
\text{s.t.} & \quad \frac{\lambda_1 30 + \lambda_2 24 + \lambda_3 96}{\lambda_1 600 + \lambda_2 300 + \lambda_3 800} \geq \eta_B \times 12 \\
& \quad \frac{\lambda_1 130 + \lambda_2 144 + \lambda_3 10}{\lambda_1 650 + \lambda_2 1200 + \lambda_3 200} \geq \eta_B \times 8 \\
& \quad \lambda_1 100 + \lambda_2 70 + \lambda_3 110 \leq 70 \\
& \quad \sum_{i} \lambda_i = 1 \\
& \quad \lambda_i \geq 0
\end{align*}
\]

(4.9a) - (4.9f)

Figure 4.1 shows the projection of the PPS convex hull on the output planes. The difference between the frontiers from NLP 4.9 and LP 4.8 becomes clear in the figure 4.2 where costs (input) for every branch are assumed to be equal. The red line is the frontier for the conventional LP 4.8 and black line is the frontier for the modified NLP.
4.9. In a two output one input situation the inputs are usually set to a common value for illustration purposes. However the ratios do not scale up or down to bring all the inputs at the same level and this should be considered in advance. Here, since all of the outputs are in the form of ratios, the two models will not change because, even in the modified version, the scaling effect will cancel out in the output constraint. Note that this is only for illustration purposes and for the true judgment of the DMUs we need to consider the input as well, and we cannot rely on the projection only. Notice that the black frontier is no longer linear. The approximation methods and algorithms used to solve the modified NLP will be discussed in chapter 7.

To see how things change, we take the reader to the detailed process for this example. The first step in both models is to estimate the PPS. The second is to define the frontier in its parts depending on the orientation. We will look into both orientations for this example. To establish the whole frontier, we select the convex hull of the points consisting of the DMUs and the hypothetically inferior DMUs following the disposability assumption. Figure 4.3 shows some points in the PPS, but we have not colored the disposable parts of PPS because it will block our way to see the other parts in 3D. However

![Diagram](image-url)
to help in the visualization of the convex-hull of feasible DMUs. Figure 4.4 sketches the boundaries of the PPS.

![Graph showing feasible DMUs generated by conventional (red) and modified (black) PPS estimators.](image)

Figure 4.3: Feasible DMUs generated by conventional (red) and modified (black) PPS estimators.

To observe the output orientation efficiency, we seek the best output for every input. In 3D it is like solving a one common input two output DEA on every plane orthogonal to the input axis. The output orientation efficient frontier consists of every plane’s boundary, it cuts through the input output planes as shown in the Figures 4.5 and 4.6. It is obvious that the frontiers (red and black) are not the same. We investigate the input orientation case for the same example. The PPS will be the same, we only look at it from a different point of view. This time we seek the best (minimum) input on the frontier for each neighborhood of \((y_1, y_2)\). Visually the results will look like figure 4.7. Again, it is clear that the convex hull in the new version (yellow) is different from the conventional DEA (red).
Figure 4.4: PPS boundaries in 3D: conventional (red), modified (black)

Figure 4.5: Forming the efficient frontier, output oriented. DEA conventional(red), DEA modified(black).
Figure 4.6: Forming the efficient frontier, output oriented. DEA conventional (red), DEA modified (black).

Figure 4.7: Frontier facets, input orientation.
4.3 Nonlinearity and Orientation

This section studies, mathematically, the presence of ratios in the input/output set with different orientations (input/output) to see if all are nonlinear. In section 4.1.2 the case of input orientation with ratios in the outputs was studied. The Output oriented case with ratios on the input side is derived in a similar manner.

\[
\begin{align*}
\max_{\lambda, \eta} & \quad \eta \\
\text{s.t.} & \quad \sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq x_{ik}, \quad i = r+1...m \quad (4.10a) \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot nx_{ij} \leq x_{ik} \sum_{j=1}^{n} \lambda_j \cdot dx_{ij}, \quad x_{ik} = nx_{ik}/dx_{ik}, \quad i = 1...r \quad (4.10b) \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot y_{ij} \geq \eta \cdot y_{ik}, \quad i = 1...s \quad (4.10c) \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \quad (4.10d) \\
& \quad \lambda_j \geq 0 \quad (4.10e)
\end{align*}
\]

For the case of Output orientation with ratios in the outputs, linearizing is not possible. Even by substituting \( \eta \cdot \lambda_j \) by \( \gamma_j \) and adding the auxiliary constraints, there is still no
guarantee that the LP 4.12 solution satisfies $\eta \cdot \lambda_j = \gamma_j$.

max $\eta_{\lambda,\eta}$\hspace{1cm}(4.11a)

s.t.

\[ \sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq x_{ik} \quad i = 1...m \] (4.11b)

\[ \sum_{j=1}^{n} \lambda_j \cdot n y_{ij} \geq y_{ik} \sum_{j=1}^{n} \eta \cdot \lambda_j \cdot d y_{ij} \quad y_{ik} = n y_{ik}/d y_{ik}, \quad i = 1...r \] (4.11c)

\[ \sum_{j=1}^{n} \lambda_j \cdot y_{ij} \geq \eta \cdot y_{ik} \quad i = r + 1...s \] (4.11d)

\[ \sum_{j=1}^{n} \lambda_j = 1 \] (4.11e)

\[ \lambda_j \geq 0 \] (4.11f)

max $\eta_{\lambda,\eta,\gamma}$\hspace{1cm}(4.12a)

s.t.

\[ \sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq x_{ik} \quad i = 1...m \] (4.12b)

\[ \sum_{j=1}^{n} \lambda_j \cdot n y_{ij} \geq y_{ik} \sum_{j=1}^{n} \gamma_j \cdot d y_{ij} \quad y_{ik} = n y_{ik}/d y_{ik}, \quad i = 1...r \] (4.12c)

\[ \sum_{j=1}^{n} \lambda_j \cdot y_{ij} \geq \eta \cdot y_{ik} \quad i = r + 1...s \] (4.12d)

\[ \sum_{j=1}^{n} \lambda_j = 1 \] (4.12e)

\[ \sum_{j=1}^{n} \gamma_j = \eta \] (4.12f)

\[ \lambda_j \geq \gamma_j \geq 0 \] (4.12g)

The same is true for Input orientation with ratios in the inputs. To summarize, when we have ratios in the opposite side of the orientation, the modified model is linear, and through linear programming, the new frontier can be assessed, but it becomes nonlinear
when ratios are present on the orientation side. However, the PPS remains the same and does not depend on the orientation. For the problems with ratios only on one side, one direction of future work of this research is to investigate if an optimal solution to the same side orientation (non-linear) can be achieved directly from an optimal solution to the opposite side orientation (linear).
Chapter 5

Comparing the Conventional and the Modified DEA Solver

In this chapter the soundness of our modified model in the presence of ratios is investigated. One profitability model based on ratios and one based on absolute values (non-ratios) are developed and tested on data from one of the major Canadian banks on their branches from coast to coast. The results from conventional and modified solvers on the ratio model and the non-ratio model are compared.

5.1 Use of DEA in Banking Industry

Banks have aggressively sought to improve their performance in many aspects of their business, including improving cash management and marketing new services that attract additional funds. In the financial services industry it is usual to measure bank performance using financial ratios, but commonly used performance ratios fail to properly consider the effects of multiple outputs and multiple inputs. For management to identify and develop ways to improve branch performance, other bank management tools that compensate for the weaknesses in accounting ratios are needed. DEA is one approach to help measure bank branch productivity and can provide insights into branch operating
efficiencies beyond those available from accounting ratios. DEA measures the efficiency of each branch in comparison to the set of branches under investigation. Its objective is to specify the subset of relatively inefficient branches and show the scale of their inefficiency in comparison to other similar branches which have been characterized as relatively efficient. Many service organizations’ managers would describe benchmarking and best-practice analyses as basic, widely accepted concepts already used in their businesses. DEA provides an objective way to identify best practices in these service organizations and it has consistently generated new insights that could lead to substantial productivity gains that were not otherwise identifiable. Branch efficiency is assessed without explicit knowledge of the input/output relationships. The assessment of a branch as relatively inefficient implies the existence in the data-set of branches (or combinations of branches) displaying greater efficiency. Similarly, the assessment of a branch as relatively efficient implies that the data-set does not contain any branches (or combinations of branches) performing more efficiently. Consequently, in the case of relative inefficiency, it can be shown that the performance of the branch in question can be improved, although relative efficiency does not preclude the existence of more efficient branches outside the selected data set. For each inefficient branch, DEA identifies an efficient reference set. This is the set of relatively efficient branches to which the inefficient branch has been most directly compared to in calculating its efficiency rating. This facilitates the exploration of the nature of inefficiencies at a branch by indicating those relatively efficient ones against which performance comparisons can be made. To summarize:

- Each bank branch being evaluated will have an efficiency score between zero and one, the branches with efficiency of one are “best practice” units, which means that they are not necessarily fully efficient but that they are not less efficient when compared to other branches in the study.

- DEA will identify the efficient reference set which is the subset of branches against which the inefficient branch was most directly compared to in calculating its effi-
Chapter 5. Comparing the Conventional and the Modified DEA Solver

43

ciency rating. This allows a manager to locate and understand the nature of the inefficiencies present by comparing his/her branch with a select subset of more efficient branches in the study. It therefore avoids the need to investigate all branches to understand the inefficiencies present, and consequently helps allocate limited managerial resources to areas where efficiency improvements are most likely to be achieved.

- DEA provides information on alternative paths that would make an inefficient branch relatively efficient. Based on these data, management can select the most feasible and cost effective path to improve each branch’s efficiency.

5.2 Evaluation Criteria

The evaluation criteria are mainly ratios corresponding to the profitability model. There are considerable number of financial ratios which can be used for bank performance evaluation. The Bank Credit Analysis Handbook lists over 80 different ratios covering the major categories of capital, asset quality, profitability and efficiency, liquidity and funding. Some of the more common measures used at the bank level include Return on Equity, Return on Assets, Expense to Revenue Ratio, and Return on Investment. At the branch level, there are also a significant number of ratios to measure operational effectiveness. These include total dollar output per employee, operating expenses per unit of output, loan loss rates, and average transactions per employee. Nonetheless, input and output specifications in the Bank application is necessarily the result of compromises between desirable model formulations and the available data. It is important to note that there is no unique way to build these models.
5.2.1 Profitability model

The profitability model assists us in evaluating the way a branch converts its expenses into revenues. Substantial research has been done on this key issue, specifically for banks. Berger and Humphrey [Berg 97] summarized studies on profit efficiency in financial institutions and during the time since that study, perhaps hundreds of studies were done. Depending on the economic foundation assumed—cost minimization or profit maximization—alternative models have appeared in the literature, but they are all of the econometric type aiming at the calibration of cost or profit functions. These functions can then be used to assess whether a given bank (or branch) is operating at the most profitable (or least costly) level. Controllable expenses by branch management are employee expenses, occupancy expenses, branch cross charges, and other operational expenses. In the spirit of profit efficiency (PROFEFF), an econometric financial performance measure with a sophisticated non-standard Fourier-flexible form that indicates how actual financial performance compares to a theoretical best practice frontier, and the Paradi et al. [Para 09] profitability model—expenses on the input side and net interest and non-interest income on the output side—our ratio and non-ratio models, shown in table 5.1, were developed intending to minimize the costs translated into input orientation. Employee expenses are comprised of branch salaries, employee commissions, and employee benefits. Occupancy and computer expenses are aggregated together, because of a relatively a high correlation (0.87) between the two and conformed to the Bank’s method of reporting as well. Cross Charges are expenses related to outsourced services. Other Expenses are general expenses; including travel, training expenses, stationery and other miscellaneous expenses. Sundry is a collection of miscellaneous charges and is mostly negative. We considered that as input as did Paradi et al. [Para 09] by bringing them to the input side, so the numbers were converted to positive values. A few branches that had positive sundry values became negative when switched to the input side so their number was adjusted to a slightly positive number (0.1). Fixed assets are considered as an input to
## Table 5.1: Choice of input and output for the profitability model

<table>
<thead>
<tr>
<th>Input Category</th>
<th>Ratio Model</th>
<th>Non-ratio Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employee Expense</td>
<td>Non-interest Earnings</td>
<td>Non-interest Earnings</td>
</tr>
</tbody>
</table>
reward the branches for the higher return on assets. Loan losses are also on the input side because these penalize branches with higher losses for taking on “riskier” clients.

On the other hand, in the ratio model on the output side, instead of considering the net interest earnings from the bank’s main lines of businesses in Wealth Management, Home Mortgages, Consumer Lending, Consumer Deposits, Commercial Lending, and Commercial Deposits, our model validates the branch’s strategy on credit and resource allocation to those businesses, and captures how profitable the branch is based on that decision. For example, instead of looking at the revenue from commercial loans on its own, we consider that the bank has enough credit, allocated collateral, or funded or invested X amount of dollars in this business. It is not advisable to use these metrics without ratios because there is a high correlation between the resource and the revenue. For comparison purposes we have also developed a non-ratio model which shares the same inputs on the expense side but considers revenues from the bank’s main lines of businesses on the output side without engaging the resource allocation strategy. We then solve the non-ratio profitability model and our profitability model with the conventional and modified solvers and compare the results. We expect the modified solver results to be in accordance with average profitability efficiency of 75 − 80% mentioned in most studies [Akhi 03], [McNu 05], [Fuen 03].

5.3 Data and summary results

The final data set included 811 branches across Canada, as certain branches were removed because of data issues. Consequently, “non-ratio”, “ratio-conventional” and “ratio-modified” input oriented BCC models were used to evaluate the branches by comparing against each other, country wide, and against their respective regions and market sizes.
Correlation

Table 5.2 shows the correlation between input variables (on the expense side). There were some correlations (> 0.5) between all, with the least belonging to the sundry. Sundry are miscellaneous revenue charges and do not depend on any line of business revenue. “Employee expenses” and “cross charges among branches” were highly correlated (90%) but since they are logically independent expenses and are still (< 95%) we kept both of them in the model. On the output side, as seen in table 5.3, in the “ratio” model, the correlations were mostly negative and negligible in value with the average of 0.16. The only considerable correlation, −30%, was between commissions and home mortgages rate of return, which is possible because to get more commissions, the branch might lower interest rate mortgages or accept even marginal ones. On the contrary, the correlations among revenues in the “non-ratio” model were much higher with the average of 0.58. The maximum correlation of 78% were found between commercial loans and commercial deposits. This could be related to the fact that lending is a function of how much money the bank has, which depends on deposits composed of commercial deposits as well as personal banking deposits. However, incorporating the resource allocation data,
this correlation reduces to 10%, meaning that the funds allocated for commercial loans are not strongly correlated to commercial deposits. The low correlations among outputs in our modified model is a testimony to one of the model’s advantages over the non-ratio model.

5.3.2 “Non-ratio” Profitability Model, DEA solver results

The average score from the “non-ratio” profitability model including all branches and using the DEA solver, was 88% with a standard deviation of 0.14. Within all branches, the major market size average score was 86% with 0.15 standard deviation, the rural market had the highest average of 94% with the lowest standard deviation of 9% while the mid size market had an average score of 88% and 0.13 standard deviation. When the same branches are re-grouped regionally, Ontario had the highest average of 90%, then came the Atlantic region with 89%, the Prairies with 88%, B.C. with 86% and the lowest, Quebec, with an average of 82%. Standard deviations were between 0.12-0.15.

Overall 354 branches were found to be efficient, 190 in major markets, 83 in mid size markets and 81 in the rural markets. Among efficient units 178 were from Ontario, 67 from the Prairies, 44 from B.C., 38 from Quebec, and 27 from the Atlantic region. Also, 83 efficient DMUs were referenced more than 10 times while 136 were not referenced at all.

Within local markets and regions:

Here we compared each branch against its peers that belong to the same market or region to remove the effects of the technical change. Each Ontario branch, for example, was compared against only Ontario branches. As expected the scores improved, the major market average was 90%, small urban markets’ average reached 96% and the rural market 97%. Ontario had the lowest average score of 92% while the Atlantic region had the highest with 99%.
<table>
<thead>
<tr>
<th>Correlations</th>
<th>Output (Rate of Returns)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Commissions</td>
</tr>
<tr>
<td>Commissions</td>
<td>1.00</td>
</tr>
<tr>
<td>Everyday banking return rate</td>
<td>-0.03</td>
</tr>
<tr>
<td>Wealth management return rate</td>
<td>-0.15</td>
</tr>
<tr>
<td>Home mortgage return rate</td>
<td>-0.30</td>
</tr>
<tr>
<td>Consumer lending return rate</td>
<td>-0.20</td>
</tr>
<tr>
<td>Commercial deposit return rate</td>
<td>-0.15</td>
</tr>
<tr>
<td>Commercial loan return rate</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Output (Revenues)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Commissions</td>
</tr>
<tr>
<td>Commissions</td>
<td>1.00</td>
</tr>
<tr>
<td>Everyday banking revenue</td>
<td>0.70</td>
</tr>
<tr>
<td>Wealth management revenue</td>
<td>0.72</td>
</tr>
<tr>
<td>Home mortgage revenue</td>
<td>0.70</td>
</tr>
<tr>
<td>Consumer lending revenue</td>
<td>0.57</td>
</tr>
<tr>
<td>Commercial deposit revenue</td>
<td>0.57</td>
</tr>
<tr>
<td>Commercial loan revenue</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Table 5.3: Summary of the Output correlations in non-ratio and modified model
5.3.3 “Conventional Ratio” Profitability Model, DEA solver results

The average score for the “conventional ratio” profitability model including all branches and using the DEA solver, was 69% with a standard deviation of 0.24. Within all branches, the major size market average score was 67% with the same 0.24 standard deviation, the mid size market had the lowest average of 63% with the same standard deviation and the small size market was the best performer compared against the combined set with the score of 84% and 0.19 standard deviation. When the same branches were re-grouped by region, the Prairies had the highest average of 82%, then came Ontario with 69%, the Atlantic with 65%, B.C. with 64% and the lowest, Quebec, with an average of 61% and 0.20 standard deviation. The other standard deviations were about the same (0.23).

Overall 203 branches were found to be efficient, 103 in the major markets, 61 in rural and 39 in the mid size market. Among efficient units 93 were from Ontario, 60 from the Prairies, 21 from B.C., 16 from the Atlantic region, and 13 from Quebec. Also, 64 efficient DMUs were referenced more than 10 times while 63 were not referenced at all. Branch 523, Prairies-rural, and branch 408, Ontario-major urban, were referenced more than 300 times each, however the sensitivity analysis in chapter 6, clearly shows the highly referenced DMUs are not outliers.

Within local markets and regions:

This time the branches were assessed against peers in their own group and when branches in major size markets were compared to each other the average efficiency increased to 80% from 67%, the mid size market score increased 20% and reached 83%, while the small size market score improved 5% to 87%. The standard deviation more or less stayed the same.

Evaluating banks in their own region, Ontario’s average score grew 5% to 74%, Quebec
with 89% had the largest increase of 28%, the Atlantic region score improved 26% to 91%,
the Prairies reached 87% and B.C. 92%.

5.3.4 “Modified Ratio” Profitability Model, our MATLAB coded DEA solver results

We developed a MATLAB coded solver to implement our modified model, because the
conventional DEA solvers on the market do not allow for constraint alterations and
modifications. To test our code we used different problem sets of various sizes in our
program and some conventional DEA software available on the market. As shown in
section 4.2 the results were exactly the same for the non-ratio cases while for the ratio
cases the commercially available DEA software failed to provide accurate projections
when ratios were present.

The average score for the “modified ratio” profitability model including all branches
and using the MATLAB coded DEA solver, was 78% with a standard deviation of 0.23.
Within all branches, the major size market average score was 77% with the same 0.23
standard deviation, the mid size market had the lowest average of 73% with the same
standard deviation and small size market was the best performer compared against the
whole data set with the score of 88% and 0.17 standard deviation.

When the same branches were re-grouped according to their regions, the Prairies had
the highest average at 89%, then came the Atlantic region with 79%, Quebec with 77%,
Ontario with 75% and B.C. with 72%. The standard deviations varied between 0.19-0.25.

Overall, 329 branches were found to be efficient, 189 in major markets, 73 in rural
and 67 in mid size markets. Among the efficient units 124 were from Ontario, 89 from
the Prairies, 47 from Quebec, 36 from B.C., and 33 from the Atlantic region.

The lowest DMU score belonged to branch 777, one of Ontario’s major urban branches.
And, 75 DMUs were referenced more than 10 times while 113 were not referenced at all.
No branch was referenced more than 145 times.
Within local markets and regions:

This time the branches were assessed against peers in their own group and when branches in major markets were compared to each other the average efficiency increased to 83% from 77%, the mid size market score increased by 15% to 88%, while the small markets’ score improved 3% to 91%. The standard deviation stayed approximately the same.

When evaluating banks in their own region, Ontario’s average score increased 5% to 74%, Quebec’s 89% had the largest increase of 28%, the Atlantic region score improved by 26% to 91%, the Prairies reached 87% and B.C. 92%.

5.4 Comparing the Efficiency Scores

Some question the need to use ratios when we could base our judgment on the results from the non-ratio model. The ratio model has more variables in it and as a result has more flexibility to a) build a broader PPS and b) show the branch under study in the best light. What has made analysts lose interest in ratio models is the results from the conventional DEA solver. Using any existing DEA software we see that none allows for ratio considerations, and the existing commercial software might set targets for DMUs that are beyond their reach in reality, or below their full capacity, by limiting themselves to the inaccurate benchmark as seen in section 4.2.

The average efficiency score reported by the modified BCC model is 78% which is 9% higher than the average reported by the conventional BCC model and 10% lower from the average of 88% from the non-ratio profitability BCC model. Only 7% of the branches found by our model to operate under 40% efficiency, while the conventional BCC model reports nearly double this number. Also, 40% of the branches are estimated to be efficient in the modified BCC model, 43% in the non-ratio BCC model while the conventional BCC model only records 25% efficient DMUs, among 811 available branches.
Since banking in Canada is a very regulated industry and typically profitable, there are strict and sophisticated measures and close monitoring in place to maintain branch performance at a high level. As a result the average profitability efficiency score of the branches are reported between 75% to 80% by other researchers [Akhi 03], [McNu 05], [Fuen 03]. The overall average efficiency recorded from the “modified” model falls well within this range, while the “non-ratio” model results are above and “conventional” model results are below. So, profitability is overestimated in the “non-ratio” model, meaning that some possible potential is missed, and performance is underestimated in the “conventional” model meaning that the good work the branches and the managment have done was not appreciated.

The selling point of DEA is that it can measure efficiency in a much simpler way on the one hand and identify the source of inefficiencies on the other hand. So the average efficiency score should be in accordance with what management and experts expect. In some cases DEA also uncovers some hidden potential for overall efficiency improvement, up to 25%, because it considers multiple inputs and outputs simultaneously and has more degrees of freedom compared with regression based analyses.

The average efficiency at the nation-wide level from our DEA tool is in line with the general view on the banking industry because the range and proportion of efficient DMUs are relevant [Akhi 03], [McNu 05], [Fuen 03] and the output/input choices are comprehensive. The conventional DEA results, however, do not agree with management’s vision of how “profit efficient” their bank branch is on average, and may discourage them to use DEA in general, because their business expertise in the field could be undermined. A major part of the blame for the unpopularity of DEA in the business community is on the negative impression some inappropriate models have caused [Mona 03].
5.4.1 National PPS

Categorized by the market and region but still assessed in a nation wide context, as seen in figure 5.1, in the modified BCC model, the major markets, mid size markets and Ontario that has the highest number of branches, scored 73 – 77%, just slightly behind the general average. The rural market and the Prairies had the highest scores, Quebec was at the average, the Atlantic region was slightly higher, and B.C, with 73%, had the lowest average score. The conventional BCC model results were lower with a similar trend in the market categories. In the regional categories, the Prairies region again, was the best, Quebec the worst, Ontario was about average, while the Atlantic region and B.C performed below the average. In the non-ratio model, the average scores declined with market size, but the regions’ shares did not follow any pattern. Ontario had the highest score right after the Atlantic region, the Prairies were at the average, B.C below
Table 5.4: Locally compared

<table>
<thead>
<tr>
<th>Region</th>
<th>Score Modified</th>
<th>Score Conventional</th>
<th>Standard Deviation Modified</th>
<th>Standard Deviation Conventional</th>
<th>% Efficient Units Modified</th>
<th>% Efficient Units Conventional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>94%</td>
<td>91%</td>
<td>0.14</td>
<td>0.15</td>
<td>79%</td>
<td>68%</td>
</tr>
<tr>
<td>B.C</td>
<td>95%</td>
<td>92%</td>
<td>0.14</td>
<td>0.16</td>
<td>81%</td>
<td>74%</td>
</tr>
<tr>
<td>Ontario</td>
<td>80%</td>
<td>74%</td>
<td>0.23</td>
<td>0.24</td>
<td>45%</td>
<td>33%</td>
</tr>
<tr>
<td>Prairies</td>
<td>92%</td>
<td>87%</td>
<td>0.17</td>
<td>0.20</td>
<td>76%</td>
<td>58%</td>
</tr>
<tr>
<td>Quebec</td>
<td>94%</td>
<td>89%</td>
<td>0.13</td>
<td>0.16</td>
<td>73%</td>
<td>58%</td>
</tr>
<tr>
<td>Average</td>
<td>91%</td>
<td>87%</td>
<td>0.16</td>
<td>0.18</td>
<td>71%</td>
<td>58%</td>
</tr>
<tr>
<td>Major</td>
<td>83%</td>
<td>80%</td>
<td>0.22</td>
<td>0.22</td>
<td>51%</td>
<td>42%</td>
</tr>
<tr>
<td>Midsize</td>
<td>88%</td>
<td>84%</td>
<td>0.19</td>
<td>0.21</td>
<td>62%</td>
<td>45%</td>
</tr>
<tr>
<td>Rural</td>
<td>91%</td>
<td>87%</td>
<td>0.15</td>
<td>0.18</td>
<td>66%</td>
<td>54%</td>
</tr>
<tr>
<td>Average</td>
<td>87%</td>
<td>84%</td>
<td>0.19</td>
<td>0.20</td>
<td>59%</td>
<td>47%</td>
</tr>
</tbody>
</table>

the average and Quebec had the lowest score. Comparing the modified model against a non-ratio profitability model, where the outputs are comprised of revenues from different lines of business and the money locked in the businesses are not considered, on average the scores are lower, meaning that there is a potential for improvement which was missed by the non-ratio model. The modified model identifies this potential because of having more degrees of freedom. The modified model not only varies the revenues to build a frontier but also varies the resource allocation measure. Branches can benefit from these additional degrees of freedom and may achieve higher scores. For example, the prairies region’s performance, on average, was underestimated in the non-ratio model because its choices of resource allocation were not equally taken into consideration.

5.4.2 Local PPS

When the DEA models were run locally, to adjust for environmental and economic circumstances, each region was compared within such region only, the average score increased in all models. The main driver for the higher DEA scores was the decrease in differentiation in the smaller categories. The data from table 5.4 show that in all models, Ontario and the Major Urban market scores did not rise significantly compared to oth-
<table>
<thead>
<tr>
<th></th>
<th>DMU # 58</th>
<th>Hypothetical Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conventional DEA</td>
<td>Modified DEA</td>
</tr>
<tr>
<td><strong>Global Efficiency Score</strong></td>
<td>0.86</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Inputs (Expenses)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Assets/Accruals</td>
<td>104.00</td>
<td>89.75</td>
</tr>
<tr>
<td>Loan Loss Experience</td>
<td>268.00</td>
<td>19.99</td>
</tr>
<tr>
<td>Sundry</td>
<td>19.00</td>
<td>13.02</td>
</tr>
<tr>
<td>Employee Expense</td>
<td>190.00</td>
<td>163.96</td>
</tr>
<tr>
<td>Computer/Space Expenses</td>
<td>83.00</td>
<td>71.63</td>
</tr>
<tr>
<td>Other Losses</td>
<td>27.00</td>
<td>23.30</td>
</tr>
<tr>
<td>Cross Charges</td>
<td>148.00</td>
<td>94.81</td>
</tr>
<tr>
<td>Total Expenses</td>
<td>839.00</td>
<td>476.45</td>
</tr>
<tr>
<td><strong>Outputs (Revenues)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commissions</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>%Everyday Banking ROR</td>
<td>6.0596</td>
<td>6.0596</td>
</tr>
<tr>
<td>%Wealth Management ROR</td>
<td>0.9331</td>
<td>0.9331</td>
</tr>
<tr>
<td>%Home Mortgage ROR</td>
<td>1.0240</td>
<td>1.2420</td>
</tr>
<tr>
<td>%Consumer Lending ROR</td>
<td>3.2547</td>
<td>3.6308</td>
</tr>
<tr>
<td>%Commercial Deposit ROR</td>
<td>6.9638</td>
<td>6.9638</td>
</tr>
<tr>
<td>%Commercial Loans ROR</td>
<td>7.6923</td>
<td>7.6923</td>
</tr>
</tbody>
</table>

Table 5.5: Input and Output Projections for DMU 58. Note: ROR = Rate of Return

ers because of their large size (40%+ of the total branch set) and significant variability reflected by higher standard deviations. The higher scores of other regions and market sizes are either because of minimized environmental factors or smaller, less diverse sample sizes, or both. Thus it is hard to identify the exact causes for changes in each local market/region.

### 5.4.3 Targets and Projection

So far we have compared the efficiency scores from the models and did not look into projections. Since the DEA model outcome is very dependent on the PPS and the frontier defined in each model, comparing the scores without looking into actual suggested
improvements might not provide fair judgment. A branch with a 70% efficiency score in one model could be more efficient (spending less or producing more) than a branch with an 80% efficiency score in another model because of a frontier shift. The modified model claims to detect any possible savings in expenses and better usage of credit in generating revenues. We compare the three methods for a specific DMU as further investigation, hence, DMU 58 was chosen, one of the Ontario Major Urban branches, an efficient branch in the non-ratio model and above average in the other two models. The modified method forms a hypothetical DMU using the linear combination of three Ontario major urban, one Ontario small urban, one Atlantic rural and one Prairie major urban branch. The conventional DEA finds a point on the input/output frontier which is a linear combination of inputs and outputs of one Ontario major urban, one Ontario small urban, two Ontario rural, two Atlantic rural, one BC major urban, one Prairie major urban and one rural branch. Tables 5.5 and 5.6 show that the hypothetical DMU (target) built by the modified DEA model is spending less and producing more, so it is superior to the one built by the conventional DEA model, in terms of cost efficiency. Another important difference in target setting is when the conventional DEA indicates to DMU 58 to increase its consumer lending return rate by 11.56% but does not indicate how to do so. Should the branch allocate less money to consumer lending or should it lend more in that market to earn more revenue? In contrast, the modified model gives specifics to DMU 58 on how to acheive the 5.5% increase in consumer lending ROR, by taking out 47% of the credit locked in consumer lending the revenue decreases 39% but the overal return rate increases 5%. Table 5.7 details the recipe for DMU 58 to increase ROR. The non-ratio model shows branch 58 as an efficient DMU and does not offer any improvements whereas the modified model shows the potential and details the changes to reach the goal.
### Table 5.6: Distance between DMU 58 and the target

<table>
<thead>
<tr>
<th>DMU #58</th>
<th>Targeted Improvement</th>
<th>Percentage of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conventional DEA</td>
<td>Modified DEA</td>
</tr>
<tr>
<td>Global Efficiency Score</td>
<td>0.14</td>
<td>0.2</td>
</tr>
</tbody>
</table>

#### Inputs (Expenses)

<table>
<thead>
<tr>
<th></th>
<th>Conventional DEA</th>
<th>Modified DEA</th>
<th>%Change</th>
<th>Modified DEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Assets/Accruals</td>
<td>-14.25</td>
<td>-20.52</td>
<td>-13.70%</td>
<td>-19.73%</td>
</tr>
<tr>
<td>Loan Loss Experience</td>
<td>-248.01</td>
<td>-251.44</td>
<td>-92.54%</td>
<td>-93.82%</td>
</tr>
<tr>
<td>Sundry</td>
<td>-5.98</td>
<td>-5.62</td>
<td>-31.48%</td>
<td>-29.57%</td>
</tr>
<tr>
<td>Employee Expense</td>
<td>-26.04</td>
<td>-37.50</td>
<td>-13.70%</td>
<td>-19.74%</td>
</tr>
<tr>
<td>Computer/Space Expenses</td>
<td>-11.37</td>
<td>-19.31</td>
<td>-13.70%</td>
<td>-23.26%</td>
</tr>
<tr>
<td>Other Losses</td>
<td>-3.70</td>
<td>-5.33</td>
<td>-13.70%</td>
<td>-19.74%</td>
</tr>
<tr>
<td>Cross Charges</td>
<td>-53.19</td>
<td>-44.55</td>
<td>-35.94%</td>
<td>-30.10%</td>
</tr>
<tr>
<td>Total Expenses</td>
<td>-362.55</td>
<td>-384.27</td>
<td>-43.21%</td>
<td>-45.80%</td>
</tr>
</tbody>
</table>

#### Outputs (Revenues)

<table>
<thead>
<tr>
<th></th>
<th>Conventional DEA</th>
<th>Modified DEA</th>
<th>%Change</th>
<th>Modified DEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commissions</td>
<td>0</td>
<td>0</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>%Everyday Banking ROR</td>
<td>0.0000</td>
<td>0.0033</td>
<td>0.00%</td>
<td>0.05%</td>
</tr>
<tr>
<td>%Wealth Management ROR</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>%Home Mortgage ROR</td>
<td>0.2180</td>
<td>0.2148</td>
<td>21.29%</td>
<td>20.98%</td>
</tr>
<tr>
<td>%Consumer Lending ROR</td>
<td>0.3761</td>
<td>0.1611</td>
<td>11.56%</td>
<td>4.95%</td>
</tr>
<tr>
<td>%Commercial Deposit ROR</td>
<td>0.0000</td>
<td>0.3759</td>
<td>0.00%</td>
<td>5.40%</td>
</tr>
<tr>
<td>%Commercial Loans ROR</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
5.4.4 Efficient Branches

A branch with a high number of peer references is doing better than the other branches, and it is “strongly” efficient. Although on its own, highly referenced branches do not suggest that a branch is an outlier, the significant difference between the highly referenced branches and the efficient DMUs indicates that the DMU might have environmental benefits or business advantages over the others. Conversely, those branches referenced only once, or not at all, are considered “weakly” efficient. That is, they could be efficient only because there is no one else to reference them against for their given set of inputs and outputs. The top ten highly referenced DMUs are expected to be both cost and profit efficient. The top ten branches were referenced between 128 and 309 times by other branches (199 on average, with a standard deviation of 66) in the conventional model, between 89 and 194 times by other branches (120 on average, with a standard deviation of 34), and between 91 and 145 times by other branches (118 on average, with a standard deviation of 21) in the modified model. The average referenced in the modified model was lower because of the higher number of efficient DMUs. The modified model
shows more equitable use of the top ten efficient DMUs in building the hypothetical target DMUs for the inefficient DMUs rather than suggesting that a DMU is an outlier (while it is not an outlier) by referring to it more than 300 times. The top ten DMUs in the modified and conventional models were split with the majority from Ontario and one to two from the Atlantic region and the Prairies. In the non-ratio model the majority of top ten, again, were from Ontario along with one from B.C. and one from Quebec. Also, 34% of the efficient branches in the modified model, 38% in the non-ratio model, and 31% in the conventional model were not referenced. The not-referenced branches in the non-ratio model generated only 7% more revenue than the average but the revenue to expense ratios were 17% higher. In the non-ratio model these percentages were 3% and 8% and in the conventional model they generated 11% less revenue and had 13% lower revenue to expense ratio.

5.4.5 Practical improvements

The input oriented average efficiency score in the modified and non-ratio model were 79% and 88% respectively. This implies that following the suggestions from the modified or non-ratio model, 21% and 12% of the expenses respectively could be cut without any cuts on the revenue side. In practice, according to the experts in the bank [Para 09], 30 – 40% of the cost cuts should be achieved easily, another 30 – 40% is achievable but needs a good effort, while the remaining 20 – 30% usually is not achievable. Since the branches were the same in the two models, and so were the average expenses, the modified model promises 9% more savings while providing the details on how to achieve it.

Although the conventional model’s average efficiency was lower with a prospect of more savings, the realized targets in practice do not bear the promised return ratios, and the path to get to those ratios is not clear to management. We saw earlier that for certain branches the conventional DEA misses the full potential for growth because of its restricted frontier and, in this manner, the full possible savings are not shown.
As discussed in section 5.4.3, the process for improvement is shown by the modified model. Although the non-ratio model provides the direction potential improvements should take, it ignores the availability of the funds to do so. With the recent (2008-2009) economic downturn in the markets, and the resulting credit crunch, bank branches may not be able to follow the path offered by the non-ratio model. But, of course, DEA is not a predictive tool and, if used as such, it must be understood that if sudden changes occur in the economic or technical environment, any DEA projections will likely be flawed. DEA is, after all, a retrospective analytical method.
Chapter 6

Analysis and Discussion of the Results

In this chapter we will analyze the findings from Chapter 5 and prove the modified model’s credentials.

6.1 Evaluation Scope

Although the evaluation criteria of all the models were profitability, the modified model takes the funds (resources) available into account to build the most profitable benchmark seen by less-profitable units as their role model. Using the “rate of returns” on the output side brings the model close to common managerial practices and connect to the experts’ vision. The conventional model also intends to consider the credit information, but as proven earlier, it fails to exploit the information correctly and generate unrealistic and distorted benchmarks and as a result incorrect projections.

The non-ratio model does not distort the PPS, but does not make use of the available funding information and so lacks the flexibility to build a frontier (hypothetical targets) as cost efficient as the modified one, as seen in section 5.4.3. However, the non-ratio model is capable of building a more revenue efficient frontier than the modified model
but of course, at a higher cost. There are some branches that were identified efficient in the modified model, but inefficient in the non-ratio model. In this case, the return ratios could not be improved further for a lower cost but the revenues can grow or stay the same for a lower cost. Management, perhaps, would prefer higher revenues for lower costs, although return rates remain the same. In such cases we see an opportunity of using non-ratio results to impose lower and higher limits on cost and revenues wherever necessary.

6.2 National Efficiency Score of Branches in Regions

DEA measures relative efficiency and in the literature there is doubt that the comparison of efficiency scores from different models without considering the actual projections is fair. To accommodate this valid concern we evaluate the soundness of the modified

![Average Input Levels of Regions Branches](image)

Figure 6.1: Average Input Levels of Regions Branches
model score by looking into an average representative of each region. No matter which model was chosen or what was the evaluation criteria we can calculate the average inputs and outputs for each region and compare them to each other. In relative comparison, Quebec was shown as a poor performer in the Non-ratio and Conventional models while the Modified model shows Quebec as an average performer. A closer look at the average expenses and revenues of all branches in Quebec reveals that, in reality, Quebec is not the lowest performer and both the expense and revenue graphs 6.1, 6.2 show Quebec place in the middle relative to other regions. The modified model is the only one that has captured this fact.

Figure 6.2: Average Output Levels of Regions Branches

In all models B.C performed below average and the Prairies region above average. The Atlantic region was at the average in the non-ratio model, above average in the modified model and below average in the conventional model. Again a closer look into the expense
and revenue graphs 6.1, 6.2 makes it clear that the Atlantic region, on average, is the most cost efficient on the input side except for its moderate fixed assets and second best in bad loans. On the Revenue side, although the absolute revenues (except for consumer lending) are lower than in the other regions, the rate of returns, if not better are not worse either. If a region uses less resources and produces less revenues it does not mean that the region is a low performer, it is doing well according to its resources. The modified model, again, is the only one that has shown the Atlantic region’s cost effective practices in generating outputs as well as some other regions and has listed it as a higher than average performer. If the Atlantic region went unnoticed in the “non-ratio” model, it was because the resource allocation information was not available to that model.

Ontario was above the average in the non-ratio model, below the average in the modified model and at the average in the conventional model. Again, according to general average trends, Ontario has a better than average record in cost saving but a distinctly lower than average outcome on the revenue side. We believe that the slightly lower than average score from the modified model reflects Ontario’s performance relative to other regions.

### 6.3 Benchmark Branches

Notable areas of the superior performance of the top ten and the average in the modified model is reported in Figure 6.1. The much lower bad loans of the top ten branches identified by our modified model is another evidence that the evaluation criteria of this model is rational. The recent (2008-2009) world-wide credit crunch in the financial markets proved that fewer bad loans is an important indicator of a bank’s financial health and business soundness. The Expense and Revenue Figures 6.3 and 6.4 show the better usage of resources of the top ten branches in the modified model. The top branches in the modified model generated 17,105 revenues out of 695,190 resources dedicated to the
Chapter 6. Analysis and Discussion of the Results

Figure 6.3: Average expenses of top ten branches

<table>
<thead>
<tr>
<th>Area</th>
<th>Difference from average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net-Non Performing Loans</td>
<td>92% less</td>
</tr>
<tr>
<td>All expenses</td>
<td>at least 50% less</td>
</tr>
<tr>
<td>Loan Losses</td>
<td>86% less</td>
</tr>
<tr>
<td>Fixed Asset</td>
<td>60% less</td>
</tr>
</tbody>
</table>

Table 6.1: Areas differentiating the top ten efficient branches from average
Figure 6.4: Average rate of returns of top ten branches
business and spent 5,194 less compared to other branches. The top ten in the non-ratio model used 2.42 times of the resources and spent 1.75 times more while generating 2.24 times more revenue.

6.4 Actual cost savings

The non-ratio model captures many more opportunities for cost savings with the focus on revenues. The modified model captures the potential to minimize expenses, with the focus on the rate of return on the revenue side when the nature of that business requires resources other than expenses (for example enough funds for lending). The conventional model’s goal is similar to the modified one but because of the mathematically incorrect formulation (for ratios) it is unable to achieve that goal.

Given average branch expenses of about a million dollars, the 21% potential for efficiency improvement by minimizing costs, in the modified model the bank can save over 200 million dollars annually. The realized actual savings will certainly be lower but still notable, especially over such a large branch network. As discussed in Chapter 5, experts believe about 70–80% of projected improvements could be realized with reasonable effort.

This means a bank with 811 branches can reduce expenses by the significant amount of 136 million dollars in the modified model. The non-ratio model, however, offers close to 78 million dollars of savings in practice, it has missed about 58 million dollars in savings that is practically achievable.

6.5 Sensitivity Analysis

Sensitivity analyses were performed for all three methods over all market sizes and regions to study the effects of eliminating one variable at a time. The top ten most referenced branches were also removed to measure their effect on the model. Removal of the highly referenced branches resulted in higher number of efficient branches, a 4% increase in
the modified method, 7% in the non-ratio method, and 1% in the conventional method. The average efficiency scores increased by 3%. The fact that changes are not dramatic, implies that outliers are not in the sample.

On the output side, the non-interest income seemed to have the highest impact on the average scores in the modified and conventional methods. This could be caused by its different nature from other outputs. While other outputs are in the form of rate of returns in these two methods, commissions are reported in the absolute terms. In the non-ratio method, where all the outputs are absolute values, the impact is not significant. Removal of any variable had caused the average scores to decrease because, branches lose one of their measures and it becomes harder to show themselves better. On the input side the removal of variables resulted in a 1 to 4% decrease in the overall efficiency scores as seen in table 6.2.
6.6 Comparing the results with the Bank’s own metrics

![Graph showing correlation between Modified scores and Bank’s Internal Metrics](image)

**Figure 6.5: Correlation between Modified scores and Bank’s Internal Metrics**

The two metrics relevant to profitability used by the banks were the revenue to expense ratio and the profit ratio. The relationship between our efficiency scores and the bank’s internal metrics was studied. It must be emphasized that a single profitability ratio cannot fully reflect the actual performance of the branches, which is a result of the complex interactions between the various parameters beyond that single measure. There was 23% and 19% correlation between the modified results and the profit ratios of the bank, which shows the results of two methods are related but do not provide the same differentiation among branches.
Chapter 7

Approximating the Nonlinear Frontier

In previous chapters it was pointed out that whenever the ratio variables are on the side of the model’s orientation the LP becomes non-linear. In the BCC case there is no way to linearize the formulation to our knowledge and we are left with approximation methods only. In the models where there is no orientation, however we were able to linearize the LP process, which is a significant step towards tackling the ratio problem. This Chapter gives a brief overview of possible ways to deal with the non-linearity issue. The actual implementation and realizations of these ideas is considered as future work in this research.

7.1 Maximum Slack Model

Building the frontier can be achieved without solving the DEA model, if frontier DMUs can be identified by other means. For example Desheng Wu et al.[Wu 05] showed that the DMU having the highest aggregated “output/input ratio” lies on the CCR frontier, while Chen and Ali [Chen 02b] proved that DMUs having the highest ratio of “aggregated output to the aggregated input” for any subset of input or outputs is a CCR frontier.
We know that the input or output orientation does not affect the PPS nor the frontier, which is the convex hull of the PPS. However, the way the DMUs’ efficiency is measured and projected, changes according to the model orientation. Knowing this fact can help us in a sense that whenever we have ratios only on the side of the model’s orientation we can build linear models for other orientation versions, or some other slack maximization model (variation of the SBM model) and solve these models to get the hypothetical DMUs on the frontier. Those hypothetical DMUs are part of the PPS, are on the surface of the convex hull, but were missed by the conventional DEA. Now we add those data points to our actual DMU list and run a conventional DEA model with the desired orientation regarding ratios as absolute values, and, if the number of DMUs is large enough, the approximation of the frontier would be close to a real one. Whenever the ratios exist on one side, Input only or Output only, we provided the solution to the case of opposite orientation in the previous chapters. It could be useful to be able to identify the PPS and the frontier for the non-linear case through the linear case, but there is no guarantee that results of the PPS from the opposite orientation leads to the points close to the real frontier in the desired orientation. The reason is that the frontier has different facets and it might occur that the identified frontier has no common edge with our desired facet. Finding how to maximize the chance of affecting the desired facet could lead us to a solution.

### 7.2 Additive model for DEA with ratio variables

The two phase BCC model required us to distinguish between input orientation or output orientation and in the presence of ratios on the side of the orientation we faced the non-linearity issue. Here we introduce an additive (ADD) model which combines both orientations in a single model and escapes the non-linearity. The general form when
having ratios both in inputs and outputs can be written as the LP in 7.1.

\[
\begin{align*}
\max \sum_{i=1}^{m} s_i^- + \sum_{i=1}^{s} s_i^+ & \\
\text{s.t.} & \\
\sum_{j=1}^{n} \lambda_j \cdot nx_{ij} - x_{ik} + s_i^- = 0 & x_{ik} = nx_{ik}/dx_{ik} \forall i = 1..q \quad (7.1a) \\
\sum_{j=1}^{n} \lambda_j \cdot dx_{ij} - s_i^- = 0 & i = q + 1..m \quad (7.1b) \\
\sum_{j=1}^{n} \lambda_j \cdot y_{ij} - y_{ik} - s_i^+ = 0 & \forall i = r + 1..s \quad (7.1c) \\
\sum_{j=1}^{n} \lambda_j \cdot ny_{ij} - y_{ik} - s_i^+ = 0 & y_{ik} = ny_{ik}/dy_{ik} \forall i = 1..r \quad (7.1d) \\
\sum_{j=1}^{n} \lambda_j = 1 & \quad (7.1e) \\
\lambda_j \geq 0 & \quad (7.1f) \\
\end{align*}
\]

The non-linear constraints 7.1e and 7.1b are linearized in the following fashion.

1.1. \[ \sum_{j=1}^{n} \lambda_j \cdot (ny_{ij} - dy_{ij} \cdot y_{ik}) - \sum_{j=1}^{n} \lambda_j \cdot dy_{ij} \cdot s_i^+ = 0 \quad \forall i = 1..r \]

\[ \omega_{ij} = ny_{ij} - dy_{ij} \cdot y_{ik} \quad \lambda_j \cdot s_i^+ = \Delta_{ij} \]

1.2. \[ \sum_{j=1}^{n} \lambda_j \cdot w_{ij} - \sum_{j=1}^{n} \Delta_{ij} \cdot dy_{ij} = 0 \quad \forall i = 1..r \]

2.1. \[ \sum_{j=1}^{n} \lambda_j \cdot (nx_{ij} - dx_{ij} \cdot x_{ik}) + \sum_{j=1}^{n} \lambda_j \cdot dx_{ij} \cdot s_i^- = 0 \quad \forall i = 1..q \]

\[ \sigma_{ij} = nx_{ij} - dx_{ij} \cdot x_{ik} \quad \lambda_j \cdot s_i^- = \Phi_{ij} \]

2.2. \[ \sum_{j=1}^{n} \lambda_j \cdot \sigma_{ij} + \sum_{j=1}^{n} \Phi_{ij} \cdot dx_{ij} = 0 \quad \forall i = 1..q \]

**Definition** DMU\(_k\) is ADD efficient if and only if \(s_i^+ = 0, S_i^- = 0, \Phi_{ij} = 0, \) and \(\Delta_{ij} = 0\)

**Theorem** DMU\(_k\) is ADD efficient if and only if it is BCC efficient.

Proof: The additive and BCC models share the same PPS, and the additive model is exactly the same as the second phase of the BCC model for a Farrel efficient unit, so if a
unit proves to be BCC efficient it is Additive efficient as well. If a unit is ADD efficient
then it is somewhere on the frontier, its inputs cannot be reduced more or outputs cannot
be increased more without worsening the situation. It means that the slacks are zero and
the variables should remain the same, it is like being multiplied by “one” which, if done,
result is the same score as BCC efficiency score.

The linearized additive model for mixed ratio and non-ratio variables in both input
and output would be as written as the LP in 7.2

\[
\begin{align*}
\max & \sum_{j=1}^{n} \sum_{i=1}^{q} \phi_{ij} + \sum_{i=q+1}^{m} s_i^- + \sum_{j=1}^{n} \sum_{i=1}^{r} \Delta_{ij} + \sum_{i=r+1}^{s} s_i^+ \quad (7.2a) \\
\text{s.t.} & \\
\sum_{j=1}^{n} \lambda_j \cdot x_{ij} - x_{ik} + s_i^- = 0 & \forall i = q+1..m \quad (7.2b) \\
\sum_{j=1}^{n} \lambda_j \cdot \sigma_{ij} + \sum_{j=1}^{n} \phi_{ij} \cdot d x_{ij} = 0 & \forall i = 1..q \quad (7.2c) \\
\sum_{j=1}^{n} \lambda_j \cdot y_{ij} - y_{ik} - s_i^+ = 0 & \forall i = r+1..s \quad (7.2d) \\
\sum_{j=1}^{n} \lambda_j \cdot w_{ij} - \sum_{j=1}^{n} \Delta_{ij} \cdot d y_{ij} = 0 & \forall i = 1..r \quad (7.2e) \\
\sum_{j=1}^{n} \lambda_j = 1 \quad (7.2f) \\
\lambda_j \geq 0 \quad (7.2g)
\end{align*}
\]

The just introduced additive model is capable of detecting inefficiencies and the direction
of the possible improvement but it does not provide a single efficiency score between zero
and one. The next step would be to develop an extension of the above additive model,
such as the SBM for the non-ratio DEA.
7.3 Numerical methods

The idea to generate a PPS numerically might appear to be unusual at first glance because in the conventional DEA the PPS is successfully enveloped by the piecewise linear frontier and generating data points in the PPS is of no use since the generated points do not contribute to the formation of the frontier. But in the case of the presence of ratios if we generate hypothetical DMUs which belong to the true PPS, we have a better estimate of the true frontier. The BCC rule for generating the PPS is that every linear convex combination of the DMUs is a new feasible DMU. We are careful in calculating the ratio inputs and outputs of such composite DMUs in the correct way as shown in section 4.1. It appears that the only information we need is a matrix containing all possible sets of lambdas (weights). As easy and simple as this may sound, to exhaust the PPS is a problem with no easy solution. Here we are dealing with a Nondeterministic Polynomial (NP) hard problem.

7.3.1 “for” Loops

In the BCC model the lambdas (weights) should sum up to one and lambdas are real values between zero and one. The first step to generate lambda sets is to limit the choices since there are infinite real numbers between zero and one. One idea is to digitize the interval, with a defined resolution. Theoretically, when we have a finite set of numbers for each lambda, we should be able to produce all combinations under the convexity constraint.

Approach 1

1. Pick a resolution: $0 \leq res \leq 1$

2. Select $\lambda_1$ from pool=$\{p|p = C \cdot res, \ 0 \leq p \leq 1, \ C \in INT\}$

3. Select $\lambda_2$ from pool=$\{p|p = C \cdot res, \ 0 \leq p \leq 1 - \lambda_1, \ C \in INT\}$
4. Select $\lambda_j$ from pool=$\{p|p = C \cdot res, \ 0 \leq p \leq 1 - \sum_{i=1}^{j-1} \lambda_i, \ C \in INT\}$, for $j = 3..n-1$

5. $\lambda_n = 1 - \sum_{i=1}^{n-1} \lambda_i$

Implementing approach 1 requires $n-1$ nested “for” loops. The number of iterations grows exponentially when the number of DMUs “$n$” increases, and MATLAB or any other programming software has performance problems with that. For instance for $res = 0.1$ and $n = 10$ there are 92,378 iterations, double the $n$ and the iterations will exceed 200 times 92,378 and reach 20,030,010.

**Theorem 2:** For $n$ DMUs and a defined resolution, $res$, there are $C(n + \frac{1}{res} - 1, \frac{1}{res})$ Lambda set solutions. Considering $\frac{1}{res}$ parts, each part equals to $res$, and $n$ unassigned $\lambda$s that can take nothing, one, or a number of those parts, the problem is like the classic question of placing $n$ unlabeled balls into $k$ labeled bins: $C(n + k - 1, n)$

Sometimes it is not essential to sweep all the possibilities in the “for” loops, random selection for $\lambda_j$ from $\{0, res, 2res, 3res, ... 1 - \sum_{i=1}^{j-1} \lambda_i\}$ reduces the number of iterations.

**Theorem 3:** $E[DMUs \ with \ nonzero \ weights(\lambda)s]=x$ if $res = 2^{-x}$

Proof: On average at each iteration we break the interval into half, so to get $n$ nonzero DMU weight on average, requires a $2^{-n}$ resolution.

For example, if the resolution is 0.001 it is expected that 10 DMUs’ weights might be nonzero and the rest will be zero, because 0.001 is approximately $2^{-10}$. To have 13 non-zero $\lambda$s the resolution should be at least .0001.

The predefined resolution in approach 1 prevents the weights from taking on some specific values. To address this shortage we came up with approach 2, where lambdas can take any real number between zero and one.

**Approach 2**

1. Randomly select $\lambda_1$ from a uniform distribution between 0 and 1.

2. Select $\lambda_2$ randomly from a uniform distribution between 0 and $1 - \lambda_1$. 
3. Select $\lambda_j$ randomly from a uniform distribution between 0 and $1 - \sum_{i=1}^{j-1} \lambda_i$.

4. $\lambda_n = 1 - \sum_{i=1}^{n-1} \lambda_i$

Approach 2 gives reasonable results except for the fact that after the first few lambdas, no room will be left for the rest of the weights and as a result the $\lambda$ set will include numerous zeros. It is essential that those zero weights be spread out, thus permutation of the nonzero weights offers different DMUs a chance to parent new DMUs. MATLAB does not have an impressive algorithm for $C(n, k)$ “n choose k” so to help it we need to narrow down our parent choices. Then the focus shifts towards identifying the DMUs suitable for parenting the new ones. The method could also be enhanced by replacing the permutation by some pattern generation based on symmetry which is beyond the scope of this work.

Note: A variation to approach 1 and 2, would be assigning each $\lambda$ a value between zero and one, either randomly from the uniform distribution, or choosing from $\{0, res, 2res, ..., 1\}$ and then normalizing them to enforce convexity. The problem with this method is when the number of DMUs are large we end up with very small weights which do not generate really different composite DMUs.

### 7.4 Partial PPS Improvement

Partial m and Quantile frontier assessment methods were introduced in chapter 3. Inspired by the idea of repeating a procedure to increase the chance of getting the right frontier we developed the Partial PPS improvement method. Generating weights for a small number of DMUs is NP hard but for a small number of DMUs it is not as big a computational problem as stated in previous sections. Hence, we generate hypothetical DMUs from a small group of randomly chosen DMUs and add the dominant ones to the set. If this procedure is repeated enough times it will lead to the PPS points. Here is a Monte-Carlo program:
1. Select randomly \( p \) DMUs from the data set

2. Generate hypothetical DMUs out of \( p \) parents

3. Run a local DEA for \( p \) parents and their children

4. Select the ones on the local DEA frontier, add them to the data set

5. Discard all the other hypothetical ones.

6. Increase Counter (limit is M large)

7. If counter is below limit, go to step 1

Since enumerating all the numbers is not possible and the interest is on the boundary of all possible combinations, we need to draw smart algorithms that are able to approximate the boundary. The procedure could be improved further through smarter approaches such as narrowing down the choices, and the pool of selection of the dominant DMUs.

1. Solve the DEA model traditionally, \( p = \) number of DMUs on the frontier.

2. Randomly choose \( \alpha > \frac{1}{2} \cdot p \) of the DMUs on the frontier and \( (1 - \alpha) \cdot p \) from the inefficient DMUs,

3. Generate hypothetical DMUs (children) out of \( p \) parents.

4. Run a local DEA for the \( p \) parents and their children.

5. Select the ones on the local DEA frontier, add them to the data set.

6. Discard all the other hypothetical ones.

7. Go to step 1, unless the convergence criteria is met. Convergence criteria is met if no new data point (a point beyond the previous known boundary) is generated, in “\( q \)” consecutive runs, here “\( q = 5 \)”.
More work should be done in this regard, seeking the rate of convergence, optimal “α”, “m” and “q” values, and the practicality of the suggested algorithms.
Chapter 8

Conclusions

This chapter gives a summary of the research, contribution of this work and directions for future research opportunities.

8.1 Summary and Contribution

We studied the problem of frontier distortion caused by the presence of ratios as variables in the conventional BCC DEA model. The advantages of working with ratios in DEA to make a connection to ratio analysis and appeal to the senior management was discussed and the inaccurate DEA models with their unrealistic results were suggested to be part of the cause of the unpopularity of DEA with respect to ratio analysis among business people.

8.2 Ratio variables on the opposite side of orientation

Various proposed models were studied and one modified model was selected due to its merits and satisfying the convexity axiom. A new model designed to maximize the slack
in the presence of ratios was proposed as a second stage to the existing modified model. A two phase BCC model able to handle ratios on the opposite side of the orientation was developed and tested on a large network of branches from a major Canadian bank. The other contribution of this work is a MATLAB coded solver for solving problems containing ratio variables since the commercial packages in the market do not let the users modify or add such constraints as are required for ratio variables.

The final data set included 811 bank branches in five regions, categorized by three market sizes. The modified model not only considered the costs and revenues but also took the credit available for the bank and the wise resource allocation to improve the overall efficiency. A conventional and non-ratio model was also developed for comparison purposes. The conventional model proved to be an incorrect model, as expected, since it treated ratios as absolute values and generated a distorted PPS and frontier. The non-ratio model worked well with revenues and expenses but did not use the resource/credit information available. As a result it did not capture all the cost saving opportunities because of having less flexibility. However it was found to be a useful tool to generate weight restrictions in the modified model, where the return ratios remain the same but the actual revenues and expenses are higher.

Here is a summary of the results:

- **Correlation**: In the modified model, because of engaging the credit and resource availability in the revenue increases, the correlation among variables dropped by 42% on average when compared to the non-ratio model where the outputs were just gross revenues.

- **Average score**: The banking industry is an established industry which proved to be profitable as the result of strict regulations and limited competition, numerous metrics, and the most important of all the experienced managers who keep the businesses on the right approach. As a result, the literature reports the average efficiency score of the banks are typically found to be between 75% and 80%.
[Akhi 03], [McNu 05], [Fuen 03]. The modified model’s average was the only one in this range, while the conventional results underestimate the current practice in the business and the non-ratio model missed some possible cost savings.

- Goal setting: The modified model provides detailed suggestions to an inefficient unit in order for it to become an efficient DMU, whereas the conventional model does not indicate how to improve the ratios. To increase the ROR of commercial Loans, it is possible to get more credit and issue more loans for more revenue, or free some of the credit, then reuse it for more profitable lines of business. With today’s credit crunch the modified model gives the experts guidance on how to make the best allocation of the available funds to maximize their profit.

- Savings: To put the savings in perspective, the modified model captured all the cost saving opportunities and suggested a 136 million dollar saving opportunity which is 58 million more than the suggested savings from the non-ratio model.

- Judgement: Closer investigation of the market sizes and regions revealed that the modified model also was more precise about the regions, relative ranking such as Quebec and the Prairies and allowed every region to be seen in the best possible light.

8.3 Ratio variables in both input and output sides

When ratios exist in both input and output variables, or ratios exist on the side of the model orientation the non-linearity issue was studied and some solutions were presented. Another contribution of this work is the introduction of the additive model when the ratios are present in both sides. The additive model is able to handle any mix of semi positive variables in the input and output side simultaneously. This model does not contain any non-linear constraints at the end and could be solved by linear programming.
solvers.

8.4 Limitations

All the proposed models for ratio variables require the numerator and denominator information of every ratio. If the financial ratios do not provide the detailed information of how they were calculated, DEA should not be used for efficiency evaluation.

The proposed models containing ratios are units invariant but are not translation invariant and cannot handle negative data.

8.5 Future Work

Extending the additive model to an SBM form, in order to generate a single efficiency score between zero and one is one challenge to be met.

Work on the non-linearity issue and develop the ideas presented here in order to build a practical tool to approximate the PPS dominant points and the efficiency frontier is another limitation for future research.
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive Model</td>
<td>DEA model which has no orientation and measures efficiency maximizing both the input and output slacks simultaneously.</td>
</tr>
<tr>
<td>Aggregated ratio analysis model</td>
<td>A model which aims to maximize the linear combination of all possible ratios in the system for each DMU and shares the same results with CCR.</td>
</tr>
<tr>
<td>BCC or VRS Model</td>
<td>DEA model which assumes a variable returns to scale relationship between inputs and outputs.</td>
</tr>
<tr>
<td>CCR or CRS Model</td>
<td>DEA model which assumes a constant returns to scale relationship between inputs and outputs.</td>
</tr>
<tr>
<td>Convexity</td>
<td>An axiom that requires the the multipliers are summed up to one, when creating linear combination of DMUs.</td>
</tr>
<tr>
<td>Correlation</td>
<td>A measure between +1 and −1, illustrating how two variables are related, 0 means no relationship, 1 means strongly related and the sign defines the same or opposite direction of relationship.</td>
</tr>
<tr>
<td>CRS (Constant Returns to Scale)</td>
<td>A measure where a proportionate increase in inputs results in an identical proportionate increase in outputs.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DEA (Data Envelopment Analysis)</td>
<td>A non-parametric, linear programming technique used for measuring the relative efficiency of units, considering multiple inputs and outputs simultaneously.</td>
</tr>
<tr>
<td>Disposability axiom</td>
<td>An assumption that says if $\text{DMU}_i$ is feasible then any DMU that is doing worse, producing less or consuming more, can be realized too.</td>
</tr>
<tr>
<td>DMU Decision Making Unit</td>
<td>Term used to describe a unit under study such as bank branch, hospital, firm, etc.</td>
</tr>
<tr>
<td>Efficient Frontier</td>
<td>The facets and edges of the PPS, representing the most efficient units.</td>
</tr>
<tr>
<td>FDH Free Disposal Hull</td>
<td>Free Disposal Hull assumption adds to the observed production data, the unobserved production points with output levels equal to or lower than those of some observed points and more of at least one input; or with input levels equal to or higher than those of some observed points and less of at least one output.</td>
</tr>
<tr>
<td>Full Efficiency</td>
<td>Full efficiency is attained by any DMU if and only if none of its inputs or outputs can be improved without worsening some of its other inputs or outputs.</td>
</tr>
<tr>
<td>Full Relative or Technical efficacy</td>
<td>Full technical efficiency is attained by any DMU if and only if, compared to other observed DMUs, none of its inputs or outputs can be improved without worsening some of its other inputs or outputs.</td>
</tr>
<tr>
<td>Input Slack factor</td>
<td>Identifies how much one of the inputs can be reduced without changing other inputs or outputs.</td>
</tr>
<tr>
<td>Input substitution factor</td>
<td>Identifies the smallest value for one specific input among the DMUs belonging to the PPS.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Input-Oriented Model:</td>
<td>DEA model whose objective is to minimize inputs while keeping outputs constant</td>
</tr>
<tr>
<td>Modified Model</td>
<td>DEA model presented in this work which imposes convexity on the DMUs when ratio variables are involved in both phases of efficiency score and Slack maximization.</td>
</tr>
<tr>
<td>Non-Discretionary Variable</td>
<td>A variable that effects the production but is beyond management's control like the weather.</td>
</tr>
<tr>
<td>NLP Non-Linear Program</td>
<td>Optimization problem with nonlinear constraints and/or objective function</td>
</tr>
<tr>
<td>Output Slack factor</td>
<td>Identifies how much one of the outputs can be increased without changing other outputs or inputs.</td>
</tr>
<tr>
<td>Output substitution factor</td>
<td>Identifies the largest value for one specific output among the DMUs belonging to the PPS.</td>
</tr>
<tr>
<td>Output-Oriented</td>
<td>DEA model whose objective is to maximize outputs while keeping inputs constant.</td>
</tr>
<tr>
<td>Partial m Frontier</td>
<td>A method for forming the frontier that does not impose convexity on the production set and allows for noise (with zero expected values) and as a result is less sensitive to outliers.</td>
</tr>
<tr>
<td>PPS Production Possibility Set</td>
<td>Given the observed data, the set of all possible input/output combinations that could exist.</td>
</tr>
<tr>
<td>Practical Frontier</td>
<td>Extending or limiting the theoretical frontier in DEA according to experts' opinion.</td>
</tr>
<tr>
<td>Glossary Term</td>
<td>Definition</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Profitability Efficiency Model</td>
<td>DEA Model that captures the business operations of a bank branch using revenues ratios as outputs and branch expenses as inputs.</td>
</tr>
<tr>
<td>Quantile Frontier</td>
<td>A continuous version of partial m method to form the frontier and more robust to the presence of outliers.</td>
</tr>
<tr>
<td>RA Ratio Analysis</td>
<td>A technique that uses the ratio of a single output to a single input and generates a relative efficiency score by dividing the aforesaid ratio by the corresponding &quot;best performer's&quot; ratio, on this specific ratio definition.</td>
</tr>
<tr>
<td>Ray Unboundedness</td>
<td>Scaling up or down of any realized DMU generates a new feasible DMU.</td>
</tr>
<tr>
<td>Reference Group</td>
<td>Set of efficient units to which the inefficient unit has been most directly compared when calculating its efficiency rating in DEA.</td>
</tr>
<tr>
<td>ROR Rate of Return</td>
<td>A measure that shows how much will be earned for each dollar invested in a business.</td>
</tr>
<tr>
<td>Scale efficiency</td>
<td>Scale efficiency represents the failure in achieving the most productive scale size and is the difference between CRS and VRS models.</td>
</tr>
<tr>
<td>Technical change</td>
<td>It is the relative efficiency of the entity when compared to a broader or newer peer groups.</td>
</tr>
<tr>
<td>VRS Variable Returns to scale:</td>
<td>A measure where a proportionate increase in inputs does not result in an identical proportionate increase in outputs.</td>
</tr>
</tbody>
</table>
References


