A Comprehensive Study of Bank Branch Growth Potential and Growth Trends through the Development of a Unique DEA Formulation and a New Restricted DEA Model

By:

Alex E. LaPlante

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

© Copyright by Alex E. LaPlante, 2015
Abstract

Alex E. LaPlante
A Comprehensive Study of Bank Branch Growth Potential and Growth Trends through the Development of a Unique DEA Formulation and a New Restricted DEA Model
Department of Chemical Engineering and Applied Chemistry
University of Toronto
Doctor of Philosophy, 2015

In light of the recent global economic downturn, it has become very apparent how important it is for Canadian citizens to have access to a stable and competitive banking system and how essential sound banking practices are in maintaining a country’s economic health. Given the current market volatility and the continuous entry of new competitors, it has become increasingly difficult for banks to maintain the product and client growth required to remain competitive and fiscally viable. In order to cut operating costs, improve their operational efficiency, and identify underperforming areas banks have turned to performance analysis techniques. Due to the multifaceted nature of the banking industry, many of these performance analyses tend to produce conflicting results or lack the robustness required to fully realize the relationships that exist. Moreover, these analysis techniques are generally applied to contemporaneous data and neglect to consider branch growth and growth trends over an extended time period. Consequently, this study aims to exploit the relationships that exist within a bank branch to provide a comprehensive picture of bank branch growth potential using a non-contemporaneous panel data set. This was accomplished by integrating Operations Research (OR) and engineering methodologies with Data Envelopment Analysis, a leading frontier efficiency approach used for studies such as this one.

DEA has been widely used in branch performance work to evaluate intermediation, profitability and production efficiency. However, traditional modeling approaches fail to assess growth potential and customer retention; two very significant components of a bank’s economic health and market stability. To address this issue, this research integrates Operations Research (OR) techniques with DEA methodologies to develop a new formulation for more accurately modeling branch growth from one time period to another. This model was found to be successful in evaluating a branch’s growth potential and with the integration of Malmquist Index techniques, and was able to identify growth trends over an extended period of time. In addition,
this study introduces a new restricted DEA model that restricts the Most Productive Possibilities Set (MPPS) to a convex subset of non-negative growth units through the restriction of the non-negative intensity variable found in the dual form of DEA. To verify both models and test the validity of the results, several data sets were used, including one provided by one of the Big Five Canadian Banks and one from a large Turkish bank.
Acknowledgements

This PhD has been a truly life-changing experience which would not have been possible without the support and guidance that I received from my family, friends and colleagues. This journey has been one of both academic and personal growth, and I am forever grateful to all who played a part, however small, in getting me to this point.

First and foremost, I would like to express my deepest appreciation to my supervisor, Dr. J.C. Paradi for his kindness, his guidance and, of course, for his candid fatherly advice. Over the last 4 years, his mentorship has been invaluable - he has taught me the importance of nurturing all aspects of life and learning from every experience, he has allowed me the freedom to explore, and the room to grow. One could not wish for a better or more supportive supervisor and mentor and for that I am eternally grateful.

I would like to acknowledge Mette Asmild for allowing me the opportunity to study in the beautiful city of Copenhagen and for providing invaluable guidance and research advice. My time abroad was truly life changing and I cannot thank her enough. Additionally, I would like to thank Dr. Lawryshyn and Dr. Saville for continually offering valuable feedback and advice throughout my time at the University and for serving on my defense committee.

I would also like to acknowledge my CMTE family- it has been an honour to share my graduate experience with such an intelligent and talented group of individuals. In particular, I would like to thank my CMTE partner in crime, Sau Yan Lee for years’ worth of interesting conversations, baking critiques and for always lending an ear when I was venting my frustrations. I am so lucky to have gained such an amazing friend.

Beyond my CMTE colleagues, I would like to thank my brothers Liem Pham and Nicholas Asuncion for putting up with me for so many years and Jennifer Luu for her belief in my ability to do great things. I couldn’t ask for better, more supportive friends. And, of course, I would like to thank Joe Steinberg for always being so helpful; taking on the roles of stand in chef, DEA specialist and world`s best listener, and for being unwaveringly caring and loving. He is my better half and, hands down, the best dancer I know.

Last but definitely not least, I thank my wonderful parents for always believing in me and for encouraging me to follow my dreams. I would be nowhere near the person I am today without their guidance and support. I am forever grateful for their sacrifices, the opportunities
they have afforded me and the devotion they have shown to parenthood. Words cannot express how much they mean to me and how blessed I feel to be their daughter.
Executive Summary

In order to remain economically viable, Banks, regardless of their geographic location, must continually grow their customer base and increase their sale of financial products. Traditional performance analysis measures focus on evaluating the production, profitability and intermediation capabilities of bank branches but fail to consider the growth of branch networks from one time period to the next. Consequently, the main objective of this study was to develop performance analysis techniques that are capable of properly evaluating growth potential and growth trends within bank branch networks. This was achieved through the use of OR technologies effectively extending the theory and methodologies of the DEA approach. The first extension develops a new restricted DEA model that allows the Most Productive Possibilities Set (MPPS) to exclusively consist of convex combinations of efficient units with non-negative growth while the second extension provides a new DEA framework that evaluates the relative growth efficiency of bank branches.

The restricted model developed herein extends existing DEA theory through the addition of a conditional constraint on the non-negative intensity variable ($\lambda$) which removes negative growth units from the MPPS. This prevents these units from being referenced by non-efficient units, effectively eliminating their effect on the efficiency scores of other units. This restricted model produces a modified frontier free of negative growth units consequently providing non-efficient units with exclusively positive growth peers. In some specific cases this model is able to assign positive-growth peers to the units removed from the MPPS providing potential insights into how positive growth can be achieved in future. Given that banks are ultimately trying to grow their market share, ensuring that negative growth units are not viewed as best performers is crucial to the usefulness and practicality of performance analysis results.

The relative growth model incorporates branch specific data from two consecutive time periods into one DEA model in order to evaluate a branch’s ability to grow its customer base, products and funds from one year to the next. The capabilities of this model were extended through the application of Malmquist techniques which provided a measure of the relative growth efficiency trends for each branch. Furthermore, Malmquist decompositions were used to assess the change in growth producing technologies along with individual branch efficiencies over time. This new model formulation provides a comprehensive and temporal look at branch growth.
This document provides an extensive review of all literature and DEA methodologies pertinent to this work. A synopsis of the Canadian and Turkish banking industries along with the banks evaluated in this study is provided followed by the objective, justification and detailed explanation of the development each technology introduced in this study. Subsequently, these methods are verified through their application to several data sets including those provided by major Canadian and Turkish Banks.

In summation, this work improves upon and extends the concepts and DEA methodologies that currently exist in order to better assess the growth efficiencies and growth trends of a bank branch network.
# Table of Contents

Abstract .................................................................................................................. ii  
Acknowledgements ............................................................................................... iv  
Executive Summary .............................................................................................. vi  
Table of Contents ................................................................................................... viii  
List of Figures ......................................................................................................... x  
List of Tables .......................................................................................................... xi  

Chapter 1: ............................................................................................................... 1  
Introduction ........................................................................................................... 1  
  1.1 Problem Definition ......................................................................................... 3  
  1.2 Method of Approach ...................................................................................... 4  

Chapter 2: ............................................................................................................... 6  
Performance Analysis .......................................................................................... 6  
  2.1 Traditional Analysis Methods ..................................................................... 6  
  2.2 Frontier Efficiency Methods ....................................................................... 8  
  2.4 Growth ........................................................................................................ 15  

Chapter 3: ............................................................................................................... 17  
Data Envelopment Analysis (DEA) ..................................................................... 17  
  3.1 DEA: A History ......................................................................................... 17  
  3.2 DEA Theory and Mathematical Formulation ............................................ 18  
  3.3 Assumptions .............................................................................................. 19  
  3.4 Constant Returns-to-Scale (CRS) Model .................................................. 21  
  3.5 Variable Returns-To-Scale (VRS) Model ................................................ 26  
  3.6 Additional DEA Models ......................................................................... 30  
  3.7 Results from DEA ................................................................................... 32  
  3.8 Extensions to Basic DEA Models .............................................................. 34  
  3.9 Strengths and Limitations of DEA ............................................................ 44  

Chapter 4: ............................................................................................................... 46  
DEA and Banking ................................................................................................. 46  
  4.1 DEA in Bank Branch Performance Evaluation ........................................... 46  

Chapter 5: ............................................................................................................... 54  
Banking Data: Background, Collection and Treatment ....................................... 54  
  5.1 Canadian and Turkish Banking Systems ..................................................... 54  
  5.2 Data Collection ......................................................................................... 58  
  5.3 Data Treatment ......................................................................................... 60  

Chapter 6: ............................................................................................................... 66  
Restricted Growth Model ..................................................................................... 66  
  6.1 Objective .................................................................................................. 66  
  6.2 Methodology .............................................................................................. 67  
  6.3 Demonstrative Application ....................................................................... 71  
  6.4 Application To Canadian Banking Data ..................................................... 75  
  6.5 Application To Turkish Banking Data ......................................................... 77  
  6.6 Comparison of Restricted Model Results ................................................. 79
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.6 EXFA Application and Discussion</td>
<td>79</td>
</tr>
<tr>
<td>6.7 Conclusion</td>
<td>81</td>
</tr>
<tr>
<td>Chapter 7:</td>
<td>83</td>
</tr>
<tr>
<td>Growth Efficiency Model</td>
<td>83</td>
</tr>
<tr>
<td>7.1 Motivation and Objective</td>
<td>83</td>
</tr>
<tr>
<td>7.2 Methodology</td>
<td>84</td>
</tr>
<tr>
<td>7.3 Growth Trends</td>
<td>87</td>
</tr>
<tr>
<td>7.4 Detailed Application</td>
<td>88</td>
</tr>
<tr>
<td>7.5 Combined Methodology</td>
<td>101</td>
</tr>
<tr>
<td>7.6 Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>Chapter 8:</td>
<td>103</td>
</tr>
<tr>
<td>Conclusions and Recommendations</td>
<td>103</td>
</tr>
<tr>
<td>Chapter 9:</td>
<td>106</td>
</tr>
<tr>
<td>Directions for Future Work</td>
<td>106</td>
</tr>
<tr>
<td>Glossary</td>
<td>110</td>
</tr>
<tr>
<td>References</td>
<td>114</td>
</tr>
<tr>
<td>Appendix A: Data Treatment and Variable Selection</td>
<td>123</td>
</tr>
<tr>
<td>Appendix B: Data Treatment R-Scripts</td>
<td>129</td>
</tr>
<tr>
<td>Appendix C: Restricted Growth Model</td>
<td>131</td>
</tr>
<tr>
<td>Section 1: R-Scripts</td>
<td>131</td>
</tr>
<tr>
<td>Section 2: Model Results</td>
<td>145</td>
</tr>
<tr>
<td>Appendix D: Period Growth Model</td>
<td>146</td>
</tr>
<tr>
<td>Section 1: R-Scripts</td>
<td>146</td>
</tr>
<tr>
<td>Section 2: Model Results</td>
<td>149</td>
</tr>
<tr>
<td>Appendix E: Preliminary Work</td>
<td>151</td>
</tr>
</tbody>
</table>
List of Figures

Figure 3.1: Graphical Representation of the CCR Model.............................................. 23
Figure 3.2: Graphical Representation of BCC Model.................................................. 27
Figure 5.1: Within Group Sum of Squares vs. Number of Clusters............................ 64
Figure 6.1: Graphical Representation- Restricted Growth Model............................... 68
Figure 6.2: Infeasibility Graph.................................................................................... 70
Figure 6.3: Wilson Outlier Statistic Graph................................................................. 73
Figure 7.1: Consecutive Year Period Growth Analysis.............................................. 85
Figure 7.2: Characterization of DMUs- Central Crosshairs....................................... 93
Figure 7.3: Characterization of DMUs- 25% Best Performers.................................... 93
List of Tables

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Table Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Frontier Approach Taxonomy</td>
<td>9</td>
</tr>
<tr>
<td>5.1</td>
<td>Canadian Bank’s Retail and Commercial Products and Services</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>Turkish Bank’s Retail and Commercial Products and Services</td>
<td>58</td>
</tr>
<tr>
<td>6.1</td>
<td>Restricted Growth Model- Conditions for Infeasibility</td>
<td>70</td>
</tr>
<tr>
<td>6.2</td>
<td>Test Set Model Formulation</td>
<td>72</td>
</tr>
<tr>
<td>6.3</td>
<td>Test Set Results Summary</td>
<td>72</td>
</tr>
<tr>
<td>6.4</td>
<td>Test Set Results- Infeasible Units</td>
<td>73</td>
</tr>
<tr>
<td>6.5</td>
<td>Wilson Outlier Statistics</td>
<td>73</td>
</tr>
<tr>
<td>6.6</td>
<td>Canadian Banking Data Model Formulation</td>
<td>75</td>
</tr>
<tr>
<td>6.7</td>
<td>Canadian Banking Data Results Summary- Without Indices</td>
<td>75</td>
</tr>
<tr>
<td>6.8</td>
<td>Canadian Banking Data Results Summary- With Indices</td>
<td>76</td>
</tr>
<tr>
<td>6.9</td>
<td>Turkish Banking Data Model Formulation</td>
<td>77</td>
</tr>
<tr>
<td>6.10</td>
<td>Turkish Banking Data Results Summary</td>
<td>78</td>
</tr>
<tr>
<td>6.11</td>
<td>EXFA Canadian Banking Data Results Summary</td>
<td>80</td>
</tr>
<tr>
<td>6.12</td>
<td>EXFA Turkish Banking Data Results Summary</td>
<td>80</td>
</tr>
<tr>
<td>7.1</td>
<td>Growth Model</td>
<td>83</td>
</tr>
<tr>
<td>7.2</td>
<td>Final Growth Model Formulation</td>
<td>89</td>
</tr>
<tr>
<td>7.3</td>
<td>Model Comparison</td>
<td>90</td>
</tr>
<tr>
<td>7.4</td>
<td>Complete Data Set- Results Summary</td>
<td>91</td>
</tr>
<tr>
<td>7.5</td>
<td>Traditional Production Model</td>
<td>92</td>
</tr>
<tr>
<td>7.6</td>
<td>Cluster Characteristics</td>
<td>94</td>
</tr>
<tr>
<td>7.7</td>
<td>Global Analysis Cluster Results</td>
<td>95</td>
</tr>
<tr>
<td>7.8</td>
<td>Local Analysis Results</td>
<td>95</td>
</tr>
<tr>
<td>7.9</td>
<td>Adjacent Malmquist Results</td>
<td>97</td>
</tr>
</tbody>
</table>
Table 7.10: Global Malmquist Results................................................................. 97
Table 7.11: Malmquist Cluster Comparison: 2010-2011 Frontier-2011-2012 Frontier.... 99
Table 7.12: Rolling Window Analysis- Average Efficiency..................................... 99
Table 7.13: Combined Model- Canadian Banking Data.......................................... 100

Table A1.1: Preliminary Canadian Bank Variable Selection............................... 123
Table A1.2: Preliminary Turkish Bank Variable Selection................................. 124
Table A1.3: Final Canadian Bank Model............................................................ 126
Table A1.4: Final Turkish Bank Model............................................................... 126
Chapter 1: Introduction

The banking sector is one of the most crucial industries in both Canadian and global societies. It provides an avenue by which the government can exercise monetary control and supplies a means of regulating the lending and borrowing of assets. As was demonstrated by the 2007-8 economic collapse and ensuing economic difficulties, sound banking practices are essential in maintaining a country’s economic health and stability. To ensure more stable economic environments, governments and global committees formulate regulations to govern banking activities. As a result, banks operate under very similar conditions and thus require adept customer service and marketing strategies to maintain or gain market share.

Even with the rapid advancement of technology, bank branches are still the main conduit through which banks handle transactions and funds. The top five Canadian domestic banks alone possess upwards of 6150 bank branches [CBA14], while the top five Turkish domestic Banks possess approximately 5,560 branches [TBAT12]. It is through this extensive network of branches that these banks are able to service their current customers and contact potential clients. It follows that a bank’s marketability and growth capabilities are heavily reliant on their branch network as well as the individual growth potential of each branch. To successfully evaluate these criteria, banks must implement performance analysis and target setting at a branch level, rather than at an institutional level.

Due to its broad applicability, branch analysis is, in many cases, more desirable and important from a managerial standpoint than institutional level analysis. It has the ability to provide information on branch performance that may lead to a better understanding of the variables and relationships that affect a bank’s efficiency and profitability. [BERG97] Bank branches are also responsible for a large portion of the value added banking provided to customers and pose the highest operational expenses for a bank. Consequently, cost management can be more effectively performed at the branch level. Continuous improvement of branch performance, which can be achieved through the development and execution of DEA based strategies, is crucial in maintaining a competitive standing in the financial industry.

Despite the many attempts to accurately measure bank branch efficiency, the multifaceted nature of bank branches and the complexity of the services they provide have made this a difficult task. [KINS80] Financial ratios are unable to simultaneously consider all variables and
are, therefore, unable to comprehensively describe branch performance. On the other hand, traditional profitability measures do offer some desirable characteristics, but comparative analysis of branches can be misleading. Their use of averages can make the identification of and comparison to top performers very difficult. Aside from the numerous profitability measures and financial ratios, banks use frontier efficiency analysis to objectively identify best practices within their organizations. [WU06a] Amongst the frontier efficiency analyses acknowledged in literature, Data Envelopment Analysis (DEA) was found to be one of the leading approaches in the banking industry.

Since its conception, the DEA methodology has made great theoretical advances, allowing for its application to a range of real world problems including those of the banking sector. [CHAR94],[SEIF96] Without the need for prior form specification of the production function, DEA provides an estimation of the production function to which each individual Decision Making Unit’s (DMU’s) efficiency score can be compared. Furthermore, DEA offers a number of advantages over traditional parametric techniques including its ability to identify reference units for each DMU. These characteristics prove to be a very useful managerial tool as it aids in establishing potential causes and methods of improvement for the identified inefficient DMUs. [EPST89]

Consequently, DEA has become one of the most widely used approaches to measure the efficiency of financial institutions.[BERG97] There exists a considerable number of papers on the use of DEA in the banking industry; however, the majority of these focus on banks at an institutional level.[PARA12] Institutional analysis has been most frequently used to assess the effects of ownership, bank type and regulatory and environmental changes, as well as bank performance and improvement, and international comparison. Whereas DEA branch analysis is applicable to a very broad range of business objectives, including the assessment of intermediation, cost, profitability, and resource allocation efficiencies and the identification of possible sources of inefficiency.

Of the aforementioned methods, the most commonly used DEA approaches to branch efficiency analysis are the production, intermediation and profitability approaches. The production approach measures a branch’s production of transaction services (outputs) based on their capital and labour (inputs). The intermediation approach involves evaluating a branch’s ability to make loans and investments (outputs) based on the monetary assets it has gathered (inputs). The profitability approach measures a branch’s profitability based on expenses (inputs) and revenues (outputs).
While each of these approaches provides useful information pertaining to branch operations their results are based on contemporaneous data and provide no information related to trends or changes over time. To overcome this shortfall, when panel data are available, Malmquist indices are applied to measure the Total Factor Productivity (TFP) which can be decomposed into technological and efficiency change components. ([CAMA06], [ASMI07], [GAGA09]) Malmquist indices are helpful in illustrating the direction of change in productivity of a unit from one time period to another; however, they are unable to provide a measure of how efficiently the unit achieves this change. Bank branches, like most institutions, are focused on growing their client base and the amount of funds they bring in. Nonetheless, no literature in the field of frontier efficiency analysis deals with assessing the growth potential of bank branches. Consequently, the main objective of this study is to assess bank branch growth efficiency through the development and application of two distinct methodologies.

1.1 Problem Definition

Since its more formal conception in the 14\textsuperscript{th} century, the banking industry has grown to be one of the most complex industries in the world. [DAVI02] Today’s banks offer an extensive variety of products and services including simple chequing and savings accounts, mortgages and loans, retirement plans and mutual funds, commercial banking, securities market participation, international finance as well as various types of insurance. It follows that their list of competitors has grown to include insurance companies, credit unions, and lenders of last resort, government organizations and even virtual internet banks. In today’s economic conditions, a bank’s ability to remain globally competitive and continue growing its market share is heavily reliant on its ability to run efficiently and to properly identify the growth potential of the branches in its branch network which can be an onerous task if the method used lacks robustness. [PARA11] Although DEA is one of the leading performance measurement techniques used in branch performance analysis, its use of strictly contemporaneous data and its inability to capture dynamic processes has, as of yet, made it inadequate for use in branch growth analysis. In an attempt to ameliorate these issues and expand DEA theory, this research focused on developing two robust methodologies that allow for a comprehensive analysis of branch growth by means of establishing an overall growth efficiency measure and by providing more pragmatic peer groups and target objectives for inefficient DMUs.

The first methodology is an indirect approach which focuses on expanding traditional DEA models to provide a means of incorporating a priori knowledge about branch growth...
without the need for altering the input and output selection. This model ensures that the efficient frontier is restricted to convex combinations of positive growth branches, consequently providing more appropriate peer groups and target objectives to inefficient units. The second methodology provides a new framework in which non-contemporaneous data can be incorporated into one distinct DEA model allowing, the growth efficiency of branches to be measured. To conclude, the two developed methodologies were combined to create one robust model that is able to provide objective measures of growth efficiency while accounting for any a priori branch knowledge that is pertinent to assessing growth performance.

1.2 METHOD OF APPROACH

This study extends the now well accepted DEA methodology to evaluate the growth efficiency of bank branch networks. It also ensures the feasibility and validates the results of these extensions through their application to multiple real and simulated data sets. The following chapters present the complete methodology used for this research; from theoretical conception to implementation and analysis.

- **Chapter 2: Performance Analysis** - provides a thorough review of relevant literature focusing on traditional performance analysis and frontier efficiency techniques used in the banking industry.

- **Chapter 3: DEA and Related Methodologies** - presents an overview of the history and development of the DEA models. Various basic DEA models and formulations are discussed and some common extensions are introduced.

- **Chapter 4: DEA and Banking** - provides a detailed literature review of DEA bank branch performance analyses.

- **Chapter 5: Banking Data: Background, Collection and Treatment** - introduces the Canadian and Turkish banking systems as well as the banks analyzed in this study. Additionally, the data sources, collection methods and environmental indices used throughout the study are reviewed and all preliminary data analysis techniques and treatments are discussed.

- **Chapter 6: Restricted Growth Model** - presents the motivation and theoretical development of the Restricted Growth Model. The model is then applied to three data
sets; a simulated set, a Canadian Branch network and a Turkish Branch network. A discussion of the results follows.

- **Chapter 7: Growth Efficiency Model** - presents the motivation and theoretical development of the Growth Efficiency Model. A real-world application of the model is then presented and the results verified. Justification for all methodologies used and a detailed explanation of the interpretation of the results are also provided.

- **Chapter 8: Combining Methodologies** - introduces the method in which the two developed methodologies were combined and applied to the Canadian Banking dataset. Results of the application are provided and discussed.

- **Chapter 9: Conclusion and Recommendations** - concludes the dissertation; providing a brief synopsis of the completed work and an outline of the theoretical contributions of this research. Recommendations and directions for future work are also provided.

- **Appendices** - include various charts and graphs from the DEA analyses performed herein as well as all references and a glossary of terms.
Chapter 2: Performance Analysis

The banking industry is a central pillar of society, playing a vital role in both global and individual economic health. With the rapid advancement of technology, the banking sector has become increasingly competitive, requiring banks to continuously assess their management practices in order to maintain viability and growth. Traditional performance measures used by banks include financial ratios, indices and regression analyses, all of which can provide important insight and benchmarking capabilities. However, these methods have intrinsic limitations and myopic viewpoints that make them insufficient as standalone methods of analysis. [PARA11] As a more robust form of evaluation, banks use frontier efficiency analyses to objectively identify best practices within their organizations. Frontier methodologies are categorized as either parametric or non-parametric and can be stochastic or deterministic. Each methodology offers unique advantages and disadvantages that should be carefully considered when choosing a frontier analysis technique. Amongst these techniques identified in literature, Data Envelopment Analysis (DEA), a deterministic non-parametric fractional linear programming technique, was found to be one of the leading approaches in the banking industry. In this chapter, a literature review on the performance analysis techniques used in the banking sector are presented. Furthermore, this chapter introduces the concepts and measurement techniques of economic growth and Total Factor Productivity (TFP).

2.1 Traditional Analysis Methods

2.1.1 Ratios

One of the most traditional and well-established performance analysis approaches used in the banking industry are Key Performance Indicators (KPI), more commonly known as financial ratios. These ratios offer the ability to measure the relationship between two numerical parameters, providing valuable insight into various aspects of banking including profitability, liquidity, asset quality and risk management. They also offer the desirable ability to quantify the change in parameter relationships over time. [GIOK08] Due to their simplicity and ease of computation, financial ratios, such as Return on Assets (ROA), Return on Equity (ROE) and Return on Investment (ROI), are still the most popular choices in many industry settings. [FRAS09]
Despite the widespread application of ratios for performance analysis, their use in the banking sector has many limitations. As a result of their simplistic nature, ratios are not very effective in assessing complex networks that perform multiple processes. [PARA04] Ratio analysis is inept in properly comparing significantly different branches or in simultaneously evaluating multiple components of a branch. [PARA11] Ratios are only capable of considering one input and provide only one output, precluding the ability to analyze situations where multiple inputs and outputs must be considered simultaneously (i.e. most banking operations). Ratios also suffer from inadequacies in assessing the effects of economies of scale, in providing benchmarks or in estimating overall performance measures of Decision Making Units (DMUs). Moreover, aggregating the results of a branch ratio analysis may result in misleading performance indicators and does not provide an objective means of determining whether a branch is actually efficient or not. [GIOK08]

2.1.2 INDICES

Index numbers are the most widely used method of measuring changes in economic variables over time. They are formulated using weighted ratios and thus can be customized to measure changes across firms, industries and countries. Index numbers that describe a number of diverse economic aspects are regularly compiled and published. These can include indices of import and export prices, price deflators for national income aggregates, and financial indices. The most commonly used economic index is Gross Domestic Product (GDP), which refers to the market value of all officially recognised final goods and services produced within a country in a given period of time. Banks employ several indices to help evaluate their branch performance. The Market Attractiveness Index, for example, compares the local population demographic with the local competitor base to determine which branches are operating in ideal environments. [COEL05]

Some of the earliest noted price indices were the Paasche and Laspeyres indices, which made their debut in the late nineteenth century. These indices are still commonly used today. Perhaps the most widely used price index is the Tornqvist index, which uses a basic geometric average of the current price of goods relative to base period prices. [IBRD04] Indices do, however, have some inherent weaknesses including the need to determine variable weights and their tendency to conceal deficient areas due to the aggregation of data. [PARA12]
2.1.3 Regression

Regression analysis is a statistical tool that is used to investigate the relationship between variables. It is a parametric technique that requires the specification of a production function, also known as the regression function. The simplest form of regression analysis, single regression, is only capable of handling either multiple inputs and a single output or multiple outputs and a single input. More complex forms of regression analysis are capable of handling multiple inputs and outputs; however, these require the use of simultaneous regression functions. Whether utilizing single or multi-analysis, regression analysis only produces an estimate for either the input or output variables, depending on which is given. [SEN90]

Comparative studies between DEA and regression analysis ([BANK96], [BOWL85], [THAN93]) suggest that there is some agreement between the two methods, and that combining them may provide an alternative analysis opportunity. However, there are some clear advantages and disadvantages between each method. Regression analysis was found to be less prone to extreme inaccuracies at the individual DMU level. This characteristic is a result of its use of the full data set to obtain estimates, making them less sensitive to data fluctuations at the individual DMU level. Regression analysis was also found to be advantageous when setting individual maximum or minimum levels when the variables are independent of one other. [THAN93]

Conversely, DEA was found to outperform regression analysis in identifying efficiencies and inefficiencies. Moreover, it is not always easy to identify the most suitable regression formula, and thus regression analysis requires a priori hypotheses, where DEA does not. [BOWL85]

2.2 Frontier Efficiency Methods

Due to the aforementioned limitations of traditional performance measures, frontier efficiency analyses have become preferred methods of evaluating performance. These methods objectively formulate an overall efficiency score for each unit, allowing for a sophisticated means of ranking units relative to one another. When dealing with more traditional, well-informed industries, frontier efficiency methods will generally provide little qualitative information that is not already known. However, when dealing with complex service industries, such as the financial service industry, easily identifying maxima and minima is not typically possible. Frontier efficiency analysis provides these complex industries with the ability to quantify a unit’s relative efficiency by comparing it with a known set of like units. These units can then be benchmarked against their peers, permitting management to objectively identify
“best-practice” units. Moreover, these techniques allow for the identification of areas of input overuse and/or output underproduction.

Frontier efficiency methods can be divided into two categories; parametric and non-parametric techniques. In parametric methods, the functional form of the efficient frontier is pre-defined or imposed a priori. Conversely in non-parametric methods only the broad class of functions; for example all increasing convex functions; or broadly defined properties are fixed a priori. This prohibits parameterization in terms of limiting the number of parameters and allows the functional form to be empirically calculated. [MURI04] Despite these differences, parametric and non-parametric methods both offer their own advantages and disadvantages, and neither is strictly preferred over the other. [MURI04], [BERG97]

Another relevant distinction between frontier efficiency methods is between deterministic and stochastic models. Deterministic models, also termed “full frontier” models, envelope all of the observations and define technical efficiency as the distance between the observed data point and the frontier. Deterministic specification assumes all deviations from the efficient frontier are due to inefficiencies and can be modified by better managing the unit. It follows that this approach should not be used when large amounts of random error exist in the data. Stochastic frontier approaches differ in that they consider deviations from the frontier to be attributed to both inefficiencies and random noise. These methods incorporate double-sided random error into the specification of the frontier model. This allows the specification errors and uncontrollable factors to be modeled independently of the technical inefficiency component. [MURI04]

The taxonomy of methods along with the most commonly used technologies are illustrated in Table 2.1. [BOGE11] Each of these is discussed in detail in subsequent sections.

**Table 2.1: Frontier Approach Taxonomy**

<table>
<thead>
<tr>
<th></th>
<th>Deterministic</th>
<th>Stochastic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parametric</strong></td>
<td><strong>Corrected Ordinary Least Squares (COLS)</strong></td>
<td><strong>Stochastic Frontier Analysis (SFA)</strong></td>
</tr>
<tr>
<td><strong>Non-Parametric</strong></td>
<td><strong>Data Envelopment Analysis (DEA)</strong></td>
<td><strong>Stochastic Data Envelopment Analysis (SDEA)</strong></td>
</tr>
</tbody>
</table>

It should be noted that for each class, there are a large number of model variations that correspond to different assumptions about the technology, the distributions etc.

When choosing the most appropriate technology, one must carefully consider the application of the model and the dataset as well as whether flexibility in the mean structure or precision in noise separation is more important. Non-parametric models are considered more flexible in terms of production economic properties as they are able to adapt their mean
structures to the data instead of relying on arbitrary assumptions. On the other hand, stochastic methods are much more robust and effective at coping with noisy data. Intuitively, combining the two methodologies, as is done in SDEA models, would provide the ideal solution. However, these technologies are complex and require more data and strong assumptions about the noise distributions. Moreover, it is thought that the lack of stochasticity in DEA can be partially compensated for by the flexible structure and the restricted structure in SFA can be compensated for by allowing random error. It follows that DEA and SFA are generally found to be very useful techniques and thus using the more complex SDEA models may not be necessary.

Although most studies involving bank performance evaluation using frontier efficiency approaches focus on the application of one method, there are a few studies that provide a comparison between methods. The studies by Bauer et al. [BAUE93], Hasan and Hunter [HASA96], Berger and Mester [BERG97b] and Berger and Hannan [BERG98] provide comparisons of two or more parametric approaches. For the most part, these studies agreed that the average efficiencies obtained from each method were comparable and relatively consistent with normal banking conditions. Perhaps more relevant to this discussion are the studies of Ferrier and Lovell [FERR90], Eisenbeis et al. [EISE97], Resti [REST97], Bauer et al. [BAUE98] and Weill [WEIL04], which provide comparisons of parametric and non-parametric methodologies. Due to the inherent differences between the parametric and non-parametric methodologies, there were more inconsistencies found in these studies. However, the presence of these discrepancies was heavily reliant on the nature of the data and the function of the analysis.

There are five main frontier approaches identified in literature as methods to evaluate the efficiency of banks. Of these, there are three parametric methods, namely Stochastic Frontier Analysis (SFA), the Distribution-Free Approach (DFA) and the Thick Frontier Approach (TFA). The remaining two approaches, Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA), are non-parametric methodologies. Stochastic Data Envelopment Analysis (SDEA) and the semi-parametric Stochastic Non-Smooth Envelopment of Data (StoNED) approach are also discussed herein.

2.2.1 Stochastic Parametric Frontier Approaches

As with all modeling techniques, stochastic parametric methods offer some distinct advantages and disadvantages. The key advantage of these models is that they allow for random error, thus reducing the chance of misidentifying error or confusing data as an inefficiency. The biggest challenge when employing these stochastic methods is accurately separating the random
error from inefficiency. [BERG97] The chief disadvantage of the parametric approach is its inherent need for the pre-specification of the functional form. [BAUE98] The wrong form specification can lead to major inaccuracies in the resulting efficiency estimates. [GREB99] In the following sections, a brief discussion of the three main types of parametric frontier approaches are discussed; namely SFA, DFA and TFA.

2.2.1.1 STOCHASTIC FRONTIER ANALYSIS (SFA)

One of the most commonly used parametric approaches is stochastic frontier analysis (SFA), also known as the econometric frontier approach. [HAO01], [BERG97a], [KAPA94] It was first introduced in 1977 when Aigner et al. [aign77], Battese and Corra [batt77], and Meeusen and Van Den Broeck [meeu77] worked simultaneously on the new method. SFA formulates a frontier for scenarios with either one input and multiple outputs or one output and multiple inputs. SFA allows for random error and assumes it follows a symmetric distribution, usually the standard normal distribution. The inefficiencies are assumed to follow an asymmetric distribution, generally the half-normal distribution. [BERG93] The use of different distribution models permits the separation and subsequent measurement of the random error and inefficiency. SFA’s assumption of a half-normal distribution for inefficiencies is fairly inflexible and presumes that most units cluster near full efficiency. Various studies ([GREE90], [YUEN93], [BERG97c]) have found that using the truncated normal or gamma distributions instead of the half-normal distribution gives additional flexibility to the evaluation of inefficiency. However, the similarity of these alternate distributions with that of the random error makes separation of the inefficiency from the random error term much more difficult. [BERG97]

2.2.1.2 DISTRIBUTION-FREE APPROACH (DFA)

The Distribution-Free Approach (DFA) ([BERG93], [deyo94]) is similar to SFA in that it specifies the functional form of the frontier, but differs in its separation of random error and inefficiency. As long as the inefficiencies are non-negative, DFA allows them to follow nearly any distribution pattern. It assumes that the average efficiency of each unit remains constant over time while the random error averages to zero. DFA estimates each unit’s inefficiencies by calculating the difference between the average residual of the data set and the average residual of the unit. Some truncation is then performed if the random error does not average to zero. [BAUE98] DFA’s assumptions suggest that efficiency does not fluctuate with time. Should fluctuations occur over time, DFA describes the average deviation of each unit from the average best practice frontier instead of the efficiency. [BERG97]
2.2.1.3 **Thick Frontier Approach (TFA)**

The Thick Frontier Approach (TFA) ([BERG91], [CLAR96], [DEYO98]) specifies a functional form but imposes no distributional assumptions on either inefficiency or random error. TFA calculates separate cost functions for the highest and lowest cost quartiles for each class size. It then assumes that inefficiencies are represented by the difference between the highest and lowest quartiles and that the deviations within the highest and lowest quartiles represent random error. Therefore, units in the lowest average-cost quartile are assumed to have above average efficiency, forming a thick frontier. It should be noted that TFA does not provide efficiencies for individual units but overall efficiency, reducing the effects of extreme data points. [BERG97]

2.2.2 **Deterministic Non-Parametric Frontier Approaches**

Unlike parametric techniques, non-parametric frontier approaches do not require prior specification of a functional form. This mitigates potential inaccuracy issues related to the pre-assignment of the functional form required in parametric methodologies. However, as many non-parametric methods are deterministic in nature, they are not without their disadvantages. Deterministic approaches assume random error does not exist. This assumption can cause deviations in efficiency measures should a unit have associated random error. Furthermore, if this unit lies on the efficient frontier, the efficiencies of its peer units may also be affected. [BAUE98] In the following sections, a brief discussion of the main deterministic non-parametric technique, Data Envelopment Analysis (DEA), is provided along with an introduction to a specialized case of DEA termed the Free Disposal Hull (FDH).

2.2.2.1 **Data Envelopment Analysis (DEA)**

Data Envelopment Analysis (DEA), first introduced by Charnes, Cooper and Rhodes in 1978 [CHAR78] extended Farrell’s concept [FARR57] of estimating technical efficiency through the comparison of each organisational unit with the efficient Production Frontier. DEA is a non-parametric linear programming technique that is used to measure the relative efficiency of comparable Decision Making Units (DMUs). It provides an estimation of the empirical production function to which each individual DMU’s efficiency score, ranging from 0 to 1, can be compared. Furthermore, DEA offers a number of advantages over traditional parametric techniques including its ability to identify reference units for each DMU. This characteristic proves to be a very useful managerial tool as it aids in determining the potential causes and remedies for the identified inefficiencies. [EPST89] Additionally, DEA does not require prior
assumptions on the specification of the production function’s form or the distribution of the observations. A detailed discussion of the history, theory and mathematical formulation of basic DEA models are provided in Chapter 3.

2.2.2.2 Free Disposal Hull (FDH)

The Free Disposal Hull (FDH) approach ([DEPR84], [TULK93], [FRIE96], [DEKK01]) is a variation of DEA where the efficient frontier includes only the DEA vertices and the free disposal hull points interior to these vertices. [BERG97] Rather than DEA’s piecewise linear frontier, FDH constructs a stepwise frontier so that efficiency measures are affected only by actual observations. [COOP07] As a result of the FDH frontier being either congruent with or interior to the DEA frontier, the FDH methodology generally produces higher estimations of average efficiency than DEA. [TULK93]

2.2.3 Stochastic Non-Parametric and Semi-Parametric Frontier Approaches

For several years, researchers have been working to develop a methodology that bridges the gap between DEA and SFA to create a technology that offers the flexibility of non-parametric approaches while accounting for random noise. This section briefly introduces two such technologies, namely the Stochastic DEA (SDEA) and the Semi-Parametric Non-Smooth Envelopment of Data (StoNED) approach.

2.2.3.1 Stochastic DEA (SDEA)

Stochastic DEA is an extension to the traditional DEA model that allows for the possibility of random factors. [BANK88] Along with the one-sided error component to account for deviations due to inefficiencies, this methodology incorporates a symmetric two-sided error component. The model is formulated as a goal programming type linear programme that combines the traditional DEA model with minimum absolute deviation (MAD) regression. The formulation non-parametrically estimates the monotonic, concave functional form of the production function and allows for the variation in the weight on the two types of deviations making it possible to obtain estimates of the frontier ranging from the external DEA frontier to the average (MAD) regression model. This additional flexibility permits the examination of the extent to which production frontier estimates and relative efficiency evaluations are sensitive to the assumptions made about the relative importance of deviations due to inefficiency or measurement and specification error.
Although this technique provides the flexibility of a non-parametric approach and accounts for random noise, it introduces added complexity and requires several other specifications. Aside from having to specify the relative weights of the two types of deviations, the distributions for all random variables must be specified as well as all of the standard deviation values. For these reasons, SDEA is still not as widely used as standard DEA and SFA techniques.

2.2.3.2 SEMI-PARAMETRIC NON-SMOOTH ENVELOPMENT OF DATA (StoNED)

The recently published work of Kuosmanen and Kortelainen [KUOS12] introduces an alternative approach called the Semi-Parametric Non-Smooth Envelopment of Data (StoNED). This is a two stage technology that satisfies monotonicity and concavity while incorporating a homoscedastic composite error term. The first stage employs convex non-parametric least squares (CNLS) to estimate the shape of the frontier, while the second stage uses method of moments or pseudo likelihood techniques to estimate the conditional expectations of inefficiency based on this CNLS frontier. Although the StoNED approach combines the appealing features of DEA and SFA, it also shares some of their limitations including the need for a large sample size to avoid issues with dimensionality and the need to make assumptions about the composite error term. Moreover, the current technology can only deal with the homoscedastic samples and single output models.

2.2.4 BOOTSTRAPPING

Aside from frontier efficiency techniques, Bootstrapping, first introduced by Efron [EFRO79], is another commonly used benchmarking technique. The bootstrapping technique is a re-sampling technique that can be used to efficiently estimate the percentiles for small random samples. It is especially useful when inferences are to be made about a complex procedure for which theoretical results are unavailable or the sample sizes are small. Moreover, bootstrapping techniques are often used in conjunction with other modeling techniques, including those previously mentioned, to verify the efficacy of the obtained standard approximations and to improve these inferences should they be inadequate. For more information, the reader is encouraged to read [EFRO93].
2.4 GROWTH

Economists and operations researchers have been studying the concept of growth for decades, yielding numerous theories and methods of measurements. It follows that choosing the appropriate theory requires careful consideration of the viewpoint of the study, the ultimate goal of the organization/entity in question and the availability of data. This section introduces several types of economic growth along with common growth measurement techniques.

2.4.1 ECONOMIC GROWTH

Economic growth is defined as the increase in the amount of goods and services produced by an economy over time. In the context of a country’s economy it is most commonly measured as the percent rate of increase in real gross domestic products (GDP). Since the turn of the 18th century, economists have been theorizing about the underlying causes of economic growth and the factors that most affect it. This has resulted in several economic theories including the Classical Growth Theory, the Neo-Classical Growth Theory and the Endogenous or New Growth Theory. [PARK97]

The Classical Growth theory, which focused mainly on the way market economies functioned and on the dynamics of economic growth, was developed by a group of economists in the 18th and 19th centuries. The main components of this theory include the production function, technological progress, investment, and the determinants of profit, the size of labour force and the wage system. It was thought the growth process relied on the rate of technological progress and population growth. The theory suggests that technological progress will lead for a period of time but will eventually lag when a fall in profits is experienced, thus preventing further accumulation of capital. This is called the state of stagnation.

The Neo-Classical Growth Theory was established in the 1950s by Robert Solow and Trevor Swan. It introduced the notion of growth as increased stocks of capital goods. The theory is based on the assumption that countries use their resources efficiently and that capital and labour increases realize diminishing returns. From here, the Neo-Classical theory makes the following predictions: 1. an increase in capital relative to labour creates growth; 2. The growth of poor countries with less capital per capita is faster because each investment produces a higher rate of return than in a country with ample capital; 3. Economies will eventually reach a point, referred to as the steady state, where any increase in capital does not produce an increase in growth. Combining these predictions with the notion that new technology allows production with fewer resources and improves the steady state level of capital, the Neo-Classical theory explains
how saving, investment and economic growth respond to population growth and technological changes.

The next big advancement in growth theory was in the late 1980s and early 1990s when economists Paul Romer and Robert Lucas, Jr. introduced the Endogenous or New Growth Theory. This theory included a more formal mathematical explanation of technological advancement than was provided in the Neo-Classical theory. Furthermore, the endogenous growth theory incorporated the new idea of human capital which encompasses the knowledge and skills that make workers productive. It was thought that human capital has increasing returns to scale therefore the steady state defined in neo-classical growth is never reached. The endogenous growth theory also conceded that technological change results from the choices that people make in the pursuit of greater profit and thus growth does not slow as capital accumulates because the rate of growth is dependent on the type of capital one chooses to invest.

When examining growth, whether for a country or an organization, it is important to keep in mind that there is an inherent cost associated with growth. This cost comes from the fact that resources are used to progress technology and not used at the present for further production or consumption.

Although these traditional growth theories do not explicitly fit the framework of this study, they provide the necessary basis for understanding the complexities of growth and the mechanism that may affect it.

2.4.2 Growth Accounting and Total Factor Productivity

Growth accounting, first introduced by Robert Solow in 1957 [SOLO57], is the procedure used to decompose growth into the contribution attributed to the increase of inputs (i.e. labour and capital) and the contribution which cannot be accounted for by observable changes in input usage. This unexplained portion of growth is taken to represent the technological progress, also known as the Total Factor Productivity (TFP) growth, which plays a pivotal role in economic growth. [LAYA87]

TFP cannot be measured directly, but instead is a residual which accounts for the effects in total output not attributed to inputs. In the context of DEA, TFP growth is most often measured using Malmquist indices. These are discussed in further detail in Section 3.8.8.
Chapter 3:
Data Envelopment Analysis (DEA)

This chapter presents an overview of the analysis technique, Data Envelopment Analysis (DEA), used for this bank branch study. To begin, a brief history and background of the technique is presented, followed by theory and the mathematical formulations of DEA. Basic DEA models are introduced, as well as relevant extensions to these models. The chapter concludes with a discussion of the results obtained from DEA and the strengths and limitations of this technique.

3.1 DEA: A HISTORY

Introduced in 1978, Data Envelopment Analysis extended Farrell’s concepts of technical and allocative efficiency. Inspired by Debreu [DEBR51] and Koopmans [KOOP51], Farrell [FARR57] proposed that the efficiency of a firm, termed as overall efficiency, could be separated into technical and allocative efficiency. Technical efficiency represented a firm’s ability to maximize output given a set of inputs, while allocative efficiency denoted the firm’s ability to optimally employ inputs given their price and the firm’s production capabilities. [COEL98] Farrell’s concept used simple ratios of a single input over a single output to measure efficiencies. He demonstrated his ideas under the restriction of linear homogeneity, forcing constant returns to scale and limiting the versatility of the technique. Furthermore, Farrell’s technique is a radial measure and measures technical efficiency relative to an isoquant; two characteristics that can lead to erroneous efficiency measures. [FARE78]

In an attempt to ameliorate these issues, Farrell and Fieldhouse [FARR62] attempted to circumvent the constant returns restriction by grouping data and adjusting for output levels. The method was based on an index of multiple inputs and outputs, where weights were assigned to all inputs and outputs for each unit measured. However, assigning a common weight to all of the units under consideration proved to be very difficult. [FARE85]

Subsequent to Farrell and Fieldhouse’ new methods, alternative linear production models with looser linear homogeneity restrictions were introduced. Amongst these models was the revolutionary Data Envelopment Analysis model proposed by Charnes, Coopers and Rhodes. [CHAR78] The DEA methodology proposed allowed each Decision Making Unit (DMU) to choose its own variable weights/multipliers. In turn, this resulted in each DMU looking as
favourable as possible to its peers. This characteristic allowed DEA to be used when assigning numerical values to variables proved difficult or when variables were qualitative in nature. Although originally developed for use in non-profit and governmental organizations, DEA has found applications in many areas of study. It has also been subject to numerous theoretical advances and methodological extensions. Most notable was the development of the BCC model by Banker et al. [BANK84], which allowed for variable returns to scale. Some other noteworthy extensions pertinent to this study include the use of categorical and non-discretionary variables, the constricted multiplier model and the Slack-Based Model (SBM). Further discussion of these extensions is provided in this chapter.

3.2 DEA THEORY AND MATHEMATICAL FORMULATION

DEA is a linear programming technique that defines the set of best-practice or frontier observations as those for which no other DMU or linear combination of DMUs has as much or more of every output for as much or less of every input. DEA produces a convex production possibilities set by connecting the best-practice observations with a piecewise linear frontier. In an input oriented model all best-practice DMUs defining the frontier are considered efficient and receive an efficiency score of 1. The units not on the frontier are considered inefficient and receive an efficiency score of less than 1. These scores are calculated by projecting the inefficient unit onto the efficient frontier. DEA also provides benchmarking and target setting capabilities for the inefficient units, allowing management to better recognize best practices and implement improvement strategies. [BERG97]

In order to successfully apply DEA to produce meaningful efficiency scores, there are some important data criteria that should be met. Firstly, DEA evaluates inefficient DMUs by comparing them to the best-practice observations; therefore, it is crucial that all DMUs are, in fact, comparable. It follows that all DMUs must operate in the same cultural environment. As previously noted, DEA does not account for random error and the efficient frontier is very sensitive to measurement error, thus, data should be carefully cleaned and all irregularities should be removed. [BERG97] Moreover, for a DMU to be properly evaluated it must have a complete set of data for all inputs and outputs. (See Appendix A for a detailed step-by-step description of the data treatment performed in this study.)

When performing a DEA analysis, one must also pay attention to the number of Degrees of Freedom (DOF) of the model. The DOF increases with the number of DMUs and decreases with the number of input and output variables. A general rule for the minimum number of DMUs
(n) is that it should exceed the greater of either: three times the sum of the number of inputs (m) and outputs (s) or the product of the number of inputs (m) and outputs (s). [COOP07] This rule is shown below:

\[ n \geq \max\{m \times s, 3(m + s)\} \quad (3.1) \]

Lastly, when building a DEA model, one must carefully choose variables that are appropriate for the model’s objectives and that properly represent the process that is being evaluated. Statistical analysis should also be completed to remove all redundancies within the variables.

The DEA methodology is based on the measure of a DMU’s efficiency. In general, efficiency is measured using a ratio of outputs/inputs. This can be denoted mathematically as:

\[ Efficiency \ Score \ of \ DMU_0 = \frac{\text{output}_0}{\text{input}_0} = \frac{\max \sum_{r=1}^{s} u_r y_r}{\sum_{i=1}^{m} v_i x_i} \quad (3.2) \]

Subject to:

\[ \frac{\sum_{r=1}^{s} u_r y_r}{\sum_{i=1}^{m} v_i x_i} \leq 1, \quad j = 1, \ldots, n \]
\[ u_r \geq 0, \forall \, r \]
\[ v_i \geq 0, \forall \, i \]

Where

\[ y_{rj}: \text{the quantity of the } r^{\text{th}} \text{ output for unit } j, \ r=1, 2, \ldots, s, \ j=1,2,\ldots, n \]
\[ u_r : \text{the weight associated with the } r^{\text{th}} \text{ output variable, } r=1, 2,\ldots, s \]
\[ x_{ij}: \text{the quantity of the } i^{\text{th}} \text{ input for unit } j, \ i=1, 2,\ldots, m, \ j=1,2,\ldots, n \]
\[ v_i : \text{the weight associated with the } i^{\text{th}} \text{ input variable, } i=1, 2,\ldots, m \]

The greater the ratio, the higher the efficiency; thus, efficiency improvement can be obtained by increasing the outputs or decreasing the inputs.

3.3 ASSUMPTIONS

DEA makes several assumptions about the technology and the shape of the production possibility set (PPS) of the DMUs being assessed. These assumptions make up the foundation of the DEA methodology and thus their specification should be a principal consideration when these methodologies are applied. The following section introduces these assumptions and provides a brief discussion of each.
3.3.1 Convexity

The convexity assumption states that any weighted average or convex combination of feasible production plans is also feasible. Additionally, when there are more DMUs, convex combinations cannot only be created from the original points but also from the convex combinations of those convex combinations. It follows that convexity serves the purpose of enlarging the PPS, especially when the number of DMUs is limited. Generally speaking, the convexity assumption improves discrimination when a small number of DMUs are under observation. Furthermore, there are several other motivations behind using the convexity assumption which include, but are not limited to:

1. Convexity assumptions can be mathematically convenient
2. Convexity can occur naturally in some contexts and provides a reasonable approximation in others.
3. Convexity is oftentimes an operationally convenient but harmless assumption as far as results are concerned.

The convexity assumption also has some shortfalls:

1. Due to the fact that a convex combination is essentially the summation of down-scaled production plans, convexity requires divisibility. This may not be realistically possible when all factors are considered.
2. Convexity does not account for the economies of scale and scope that exist in many industries.

3.3.2 Free Disposability (Monotonicity)

The free disposability assumption states that amounts of unnecessary inputs and amounts of unwanted outputs can be freely discarded. In order words, excess inputs can be freely disposed of. Thus, if we can produce a quantity of outputs with a given quantity of inputs, then we can produce that same quantity of outputs with more inputs. This assumption also works in reverse in that if a quantity of inputs is able to produce a quantity of output that same amount of input can also produce less output. This assumption is, in most cases of constructing an empirical reference technology, a safe and weak regularity assumption that is able to provide directly identifiable real peer units rather than mathematical combinations of units. Moreover, when the set of DMUs is adequately large the free disposability assumption is able to perform comparisons between the DMUs. However, if the set of DMUs is small the discriminatory power of the technology reduces as the majority of the firms will be identified as efficient, subsequently providing little to
no improvement insights. It should also be noted that in instances of joint production, the disposability assumption may not hold and weaker types of disposability assumptions must be used. For example, it can be assumed that inputs or outputs can be reduced proportionally.

3.3.3 **Returns to Scale**

The Returns to Scale (RTS) assumptions indicate that rescaling is possible, the extent and nature of which is dependent on which assumption is chosen. The weakest assumption is the Variable Returns to Scale assumption which states that no rescaling is possible. The strongest assumption is Constant Returns to Scale assumption where any production combination can be arbitrarily scaled up or down. Between these two lays the Non-Increasing Returns to Scale assumption that allows any degree of downscaling but restricts upscaling. In other words, it cannot be disadvantageous to be small but it may be so to be large. The last assumption is the least commonly used Non-Decreasing Returns to Scale where it can be disadvantageous to be small but not so to be large. [BOGE11]

3.3.4 **Additivity**

The additivity assumption states that the sum of feasible production plans is feasible as well. This logic essentially rules out positive or negative externalities between the summed production plans, making it an appealing assumption. However, additivity models are much more involved from a mathematical context and may require more complex programming. It is this added complexity that makes the additivity assumption the least commonly applied assumption.

3.3.5 **No Free Lunch**

This is simply states that no output can be produced without an input effectively ensuring that the PPS is realistically feasible.

3.3.6 **Minimum Extrapolation**

The minimum extrapolation principle captures the essence of DEA by ensuring that the PPS is the smallest subset of which satisfies the imposed production assumptions.

3.4 **Constant Returns-to-Scale (CRS) Model**

Introduced by Charnes, Cooper and Rhodes, the Constant Returns-to-Scale (CRS) model, also referred to as the CCR model, was the first formulated DEA model. [CHAR78] As the name implies, this model was built on the assumption that the analysed units are operating under
constant returns-to-scale meaning that, regardless of operation scale, increases in inputs result in proportional increases in outputs. The CRS model assigns weights \((u, v)\) to each input \((X_j = \{x_{ij}\})\) and output \((Y_j = \{y_{rj}\})\) variable in order to maximize each unit’s relative efficiency score \((\theta)\), with no score exceeding one. Essentially, the DEA model is attempting to make each DMU look as favorable as possible.

The CCR model provides a measure known as the overall technical efficiency, which aggregates both technical efficiency and scale efficiency. Efficient DMUs have an efficiency score of one \((1)\) with both slacks from the efficiency \((s^-\text{ and } s^+)\) being zero, while inefficient DMUs have efficiency scores that range from 0 to 1. Additionally, DMUs can be termed as weakly efficient if there are slacks present and their efficiency score is one.

The CCR model has the following fractional formulation:

\[
\begin{align*}
\text{Maximize} & \quad h_0 = \frac{u_1y_{10} + u_2y_{20} + \cdots + u_sy_{s0}}{v_1x_{10} + v_2x_{20} + \cdots + v_mx_{m0}} \\
\text{Subject to:} & \quad \frac{u_1y_{1j} + \cdots + u_sy_{sj}}{v_1x_{1j} + \cdots + v_mx_{mj}} \leq 1 \quad (j = 1, \ldots, n) \\
& \quad u_1, u_2, \ldots, u_s \geq \varepsilon \\
& \quad v_1, v_2, \ldots, v_m \geq \varepsilon \\
\text{Where:} & \quad x_{ij} \text{ is the amount of the } i^{th} \text{ input to unit } j, \ i=1, 2, \ldots, m, j=1,2,\ldots, n \\
& \quad v_i \text{ is the weight given to the } i^{th} \text{ input, } i=1, 2, \ldots, m \\
& \quad y_{rj} \text{ is the amount of the } r^{th} \text{ output from unit } j, \ r=1, 2, \ldots, s, j=1,2,\ldots, n \\
& \quad u_r \text{ is the weight given to the } r^{th} \text{ output, } r=1, 2, \ldots, s \\
& \quad \varepsilon \text{ is non-Archimedean}
\end{align*}
\]

This ratio model evaluates the relative performance of the DMUs given the observed performance of the production possibilities set. The denominator represents a weighted sum of ‘m’ inputs used by the DMU to produce a weighted sum of ‘s’ outputs represented by the numerator. Each DMU is assigned multipliers to give the DMU the highest efficiency score possible while ensuring that no other DMU’s score exceeds one. To prevent any of the assigned multipliers from being zero or negative, \(\varepsilon\), a non-Archimedean constant having a value smaller than any positive valued real number, is used as a constraint in the model. \[COOP07\]

A graphical representation of the CCR model is presented below in Figure 3.1. In the case of one input and one output, the CCR efficient frontier begins at the origin and continues as a straight line. This frontier consists solely of efficient units enveloping the remaining inefficient
DMUs. As is shown, the only efficient DMU in this data set is DMU H; hence, it is the only DMU on the frontier. The remaining DMUs are all considered inefficient.

The CCR model can either be input oriented or output oriented. Both of these CCR models are scale invariant, allowing for variable adjustment, and unit invariant, allowing variables with differing units to appear in the same model. However, neither CCR model is translation invariant as the frontier must pass through the origin. Moreover, scalar changes to the variables in either CCR model will result in a change in slope of the frontier, thus altering the efficiency scores of the DMUs. In the following sections, the input-oriented and the output-oriented CCR models are introduced using a less computationally intensive linear programming model derived from the aforementioned fractional CCR model.

![Figure 3.1: Graphical Representation of the CCR Model](image-url)

### 3.4.1 INPUT ORIENTED CRS MODEL

The input-oriented CCR model aims to minimize inputs for the same observed output. This CRS model has both a primal and a dual form. To develop the primal linear programming model (3.4), the denominator of the CCR ratio model is used as the normalizing constraint, while the numerator becomes the objective function. The dual linear programming model (3.5) assigns a dual variable to each constraint in the primal model. This model is carried out in two stages; first the radial efficiency is calculated, followed by the mix efficiency.
Primal CCR Input-Oriented Model (3.4)

Maximize: \[ w_0 = \sum_{r=1}^{s} u_r y_{r0} \]

Subject to
\[ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, j = 1, \ldots, n \]
\[ \sum_{i=1}^{m} v_i x_{i0} = 1 \]
\[ -u_r \leq -\epsilon \]
\[ -v_i \leq -\epsilon \]

Dual CCR Input-Oriented Model (3.5)

Minimize: \[ z_0 = \theta - \epsilon [\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+] \]

Subject to
\[ y_{r0} = \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+, r = 1, \ldots, s \]
\[ 0 = \theta x_{i0} - \sum_{j=1}^{n} x_{ij} \lambda_j - s_i^-, i = 1, \ldots, m \]
\[ 0 \leq \lambda_j, \text{ for } j=1, \ldots, n \]
\[ 0 \leq s_i^-, \text{ for } i=1, \ldots, m \]
\[ 0 \leq s_r^+, \text{ for } r=1, \ldots, s \]

The dual model introduces some variables that were not present in the fractional form. The \( \theta \) variable represents the proportional input reduction applied to the DMU under consideration. This reduction is applied to all of the DMU’s inputs in order to bring the unit closer to the efficient frontier. This variable also denotes the required percent of inputs inefficient DMUs would be required to theoretically produce the same amount of outputs. The optimal solution \( \theta^* \) will belong to \((0,1]\), with one representing full efficiency or radial efficiency. Furthermore, the dual form employs a set of non-negative intensity variables, \( \lambda_1, \ldots, \lambda_n \), that represent the weight of each of the n DMUs.

The presence of the non-Archimedean constant ensures that the mix efficiency is calculated after the radial efficiency by assigning the slacks a lower level of importance. It follows that once the model has calculated the radial efficiency of the DMU it will then calculate the input and output slack variables, also referred to as input excesses, \( s_i^- \), and output shortfalls, \( s_r^+ \). If slacks are present, the DMU is considered to have mix inefficiencies. Thus, it will require further reductions beyond the optimal \( \theta^* \) reductions and the input proportions will need to be adjusted. If both slacks are zero and \( \theta^* \) is equal to one, then the DMU is considered fully efficient. As previously mentioned, if \( \theta^* \) is equal to one with non-zero slacks, then the DMU is considered radially efficient with mix inefficiencies, or weakly efficient.
The input-oriented CCR model provides target setting for inefficient DMUs by referring these units to the efficient frontier formed by the DMU’s unique reference set of efficient DMUs denoted as \( E_0 \). As shown in formulas (3.6) and (3.7), DMU \( 0 \) is projected to point \((\tilde{x}_0, \tilde{y}_0)\) on the frontier, where \((\tilde{x}_0, \tilde{y}_0)\) are the coordinates of a virtual linear composite used to evaluate the performance of DMU \( 0 \). Moreover, \((\tilde{x}_0, \tilde{y}_0)\) denotes the theoretically feasible efficient production that DMU \( 0 \) should strive for.

\[
\tilde{x}_{i0} = \theta^* x_{i0} - s^-_i \quad (i=1,...,m) \tag{3.6}
\]
\[
\tilde{y}_{r0} = y_{r0} + s^+_r \quad (r=1,...,s) \tag{3.7}
\]

### 3.4.2 OUTPUT ORIENTED CRS MODEL

The output-oriented CCR model aims to maximize outputs while utilizing the same amount of observed inputs. Much like the input-oriented model, the output-oriented model has both a primal and dual form as shown below:

#### Primal CCR Output-Oriented Model (3.8)

Minimize: \( \sum_{i=1}^{m} p_i x_{i0} \)

Subject to \( \sum_{r=1}^{s} q_r y_{rj} - \sum_{i=1}^{m} p_i x_{ij} \leq 0, j = 1, ..., n \)
\( \sum_{r=1}^{s} q_r y_{r0} = 1 \)
\( -p_i \leq -\varepsilon, i = 1, ..., m \)
\( -q_r \leq -\varepsilon, r = 1, ..., s \)

#### Dual CCR Output-Oriented Model (3.9)

Maximize: \( \eta - \varepsilon [\sum_{i=1}^{m} t^-_i + \sum_{r=1}^{s} t^+_r] \)

Subject to \( x_{i0} = \sum_{j=1}^{n} x_{ij} \mu_j + t^-_i, i = 1, ..., m \)
\( 0 = \eta y_{r0} - \sum_{j=1}^{n} y_{rj} \mu_j + t^+_r, r = 1, ..., s \)
\( 0 \leq \mu_j, \text{ for } j=1,...,n \)
\( 0 \leq t^-_i, \text{ for } i=1,...,m \)
\( 0 \leq t^+_r, \text{ for } r=1,...,s \)

This model functions in the same manner as the input-oriented CRS model. In the output-oriented model the output increase applied to the DMU under consideration is represented by \( \eta \). This increase is applied to all of the DMU’s outputs in order to bring it closer to the frontier.
Similar to the $\theta$ variable, $1/\eta$’s optimal solution will belong to $(1,0]$, with one representing full efficiency or radial efficiency. If a DMU is inefficient under the output-oriented model, the $\eta$ variable represents the radial expansion in outputs the DMU could theoretically produce given the same amount of inputs. Similar to the input-oriented model, the output-oriented dual form employs non-negative intensity variables, $\mu_1, \ldots, \mu_n$, that represent the weight of each of the $n$ DMUs.

An optimal solution for the output-oriented dual model can be derived from the optimal solution of the input-oriented model using the following relations:

$$\theta = 1/\eta$$  \hspace{1cm} (3.10)

Moreover, the slacks ($t^-, t^+$) of the output oriented model can be related back to the input oriented model through its optimal solution:

$$t^-^* = s^-^*/\theta^*, \quad t^+^* = s^+^*/\theta^*$$  \hspace{1cm} (3.11)

As with the input-oriented model, DMUs in an output-oriented model are fully efficient if and only if $\eta^*$, the optimal expansion, is equal to one and all optimal slacks are zero. For target setting purposes, inefficient DMUs can be projected to the efficient frontier using the following formulation:

$$\hat{x}_{i0} = x_{i0} - t^-_{i}^* \quad (i=1,\ldots,m)$$  \hspace{1cm} (3.12)

$$\hat{y}_{r0} = \eta^* y_{r0} + t^+_{r} \quad (r=1,\ldots,s)$$  \hspace{1cm} (3.13)

### 3.5 Variable Returns-to-Scale (VRS) Model

The Variable Returns to Scale (VRS) Model or BCC model, first introduced in 1984 by Banker, Charnes and Cooper, allows more flexibility than the CCR model by providing a variable returns-to-scale DEA formulation. [BANK84] As is shown in Figure 3.2, the BCC frontier does not cross through the origin as did the CCR model, but has a piecewise linear frontier that is concave in shape. This frontier, comprised of best performing DMUs, encapsulates the inefficient DMUs.
The BCC model is formulated in much the same way as the CCR model with the exception of the addition of the $\tilde{u}_o$ variable, which accounts for economies of scale. Given that a unique optimal solution is found, which may not always be the case: $\tilde{u}_o < 0$ indicates that the units are operating under increasing returns-to-scale, $\tilde{u}_o = 0$ indicates constant returns-to-scale and $\tilde{u}_o > 0$ indicates decreasing returns-to-scale. It should also be noted that although the $\tilde{u}_o$ variable estimates the scale economies, the BCC model measures technical efficiency alone and does not account for scale inefficiencies. [COOP07]

As with the CCR model, the BCC model can be either input-oriented or output-oriented. The following sections provide an introduction to these two models, outlining their formulations and target setting techniques.

3.5.1 **INPUT ORIENTED VRS MODEL**

The formulation of the input-oriented VRS model maximizes the efficiency scores of $n$ units. The fractional VRS model is depicted below:
Maximize \[ \theta = \frac{\sum_{r=1}^{s} u_r y_{r0} - \bar{u}_0}{\sum_{i=1}^{m} v_i x_{i0}} \] (3.14)

Subject to:
\[ \frac{\sum_{r=1}^{s} u_r y_{rj} - \bar{u}_0}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \quad (j = 1, \ldots, n) \]
\[ u_r \geq \varepsilon; \quad r = 1, \ldots, s \]
\[ v_i \geq \varepsilon; \quad i = 1, \ldots, m \]
\[ \bar{u}_0: \text{free in sign} \]

Akin to the CCR model, the fractional BCC model can be transformed into less computationally intensive formulations; namely the primal and dual forms.

**Primal BCC Input-Oriented Model** (3.15)

Maximize: \[ w_0 = \sum_{r=1}^{s} u_r y_{r0} - \bar{u}_0 \]

Subject to
\[ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \bar{u}_0 \leq 0, \quad j = 1, \ldots, n \]
\[ \sum_{i=1}^{m} v_i x_{i0} = 1 \]
\[ -u_r \leq -\varepsilon \]
\[ -v_i \leq -\varepsilon \]
\[ \bar{u}_0: \text{free in sign} \]

**Dual BCC Input-Oriented Model** (3.16)

Minimize: \[ z_0 = \theta - \varepsilon [\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+] \]

Subject to
\[ y_{r0} = \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+, \quad r = 1, \ldots, s \]
\[ 0 = \theta x_{i0} - \sum_{j=1}^{n} x_{ij} \lambda_j - s_i^-, \quad i = 1, \ldots, m \]
\[ 1 = \sum \lambda_j \]
\[ 0 \leq \lambda_j, \text{ for } j=1,\ldots,n \]
\[ 0 \leq s_i^-, \text{ for } i=1,\ldots,m \]
\[ 0 \leq s_r^+, \text{ for } r=1,\ldots,s \]

Aside from the addition of the \( \bar{u}_0 \) variable, the primal BCC form functions in the same fashion as the primal CCR form. The sole difference between the CRS dual form and the BCC dual form is that the \( \lambda_j \) terms are forced to sum to one. This restriction removes the CCR model requirement that all efficient DMUs are scale efficient. Moreover, this constraint reduces the
feasible region for the linear program from a convex cone defined by the DMUs to a convex hull covering all the DMUs. This, in turn, increases the number of efficient DMUs. [CHAR94] Aside from this, the dual BCC and dual CCR formulations function in the same way.

As with the CCR model, the DMUs in the BCC model are only efficient if $\theta$ is equal to one and the slacks ($s^+_r, s^-_i$) are equal to zero. The target projection to point ($\hat{x}_0, \hat{y}_0$) on the efficient frontier can be obtained through the following formulation:

$$\hat{x}_{i0} = \theta^{\ast}_{vrs}x_{i0} - s^{-}_i \quad (i=1,...,m) \quad (3.17)$$

$$\hat{y}_{r0} = y_{r0} + s^+_r \quad (r=1,...,s) \quad (3.18)$$

### 3.5.2 OUTPUT ORIENTED VRS MODEL

The primal and dual formulations for the output-oriented BCC model are as follows:

**Primal BCC Output-Oriented Model** (3.19)

Minimize: $\sum_{i=1}^{m} p_i x_{i0} - \tilde{v}_0$

Subject to

- $\sum_{r=1}^{s} q_r y_{rj} - \sum_{i=1}^{m} p_i x_{ij} - \tilde{v}_0 \leq 0, j = 1, \ldots, n$
- $\sum_{r=1}^{s} q_r y_{r0} = 1$
- $-p_i \leq -\varepsilon$
- $-q_r \leq -\varepsilon$
- $\tilde{v}_0$: free in sign

**Dual BCC Output-Oriented Model** (3.20)

Maximize: $\eta - \varepsilon[\sum_{i=1}^{m} t^-_i + \sum_{r=1}^{s} t^+_r]$

Subject to

- $x_{i0} = \sum_{j=1}^{n} x_{ij}\mu_j + t^-_i, i = 1, \ldots, m$
- $0 = \eta y_{r0} - \sum_{j=1}^{n} y_{rj}\mu_j + t^+_r, r = 1, \ldots, s$
- $1 = \sum \mu_j$
- $0 \leq \mu_j$, for $j=1,\ldots,n$
- $0 \leq t^-_i$, for $i=1,\ldots,m$
- $0 \leq t^+_r$, for $r=1,\ldots,s$

Similar to the input-oriented model, the output-oriented BCC model varies slightly from the CCR model.

---

A Comprehensive Study of Bank Branch Growth Potential and Growth Trends 29
Lastly, a DMU is VRS efficient if and only if $\eta$ is equal to 1 and the slacks $(t_i^+, t_i^-)$ are equal to zero. Inefficient DMUs can target set using the following projection:

$$\hat{x}_{i0} = x_{i0} - t_i^{-*} \quad (i=1,\ldots,m) \quad (3.21)$$
$$\hat{y}_{r0} = \eta_{vrs} y_{r0} + t_r^{+*} \quad (r=1,\ldots,s) \quad (3.22)$$

### 3.6 ADDITIONAL DEA MODELS

Aside from input-oriented and output oriented models, DEA offers a range of non-oriented models. These models provide the capability of simultaneously maximizing outputs while minimizing inputs. Non-oriented models have the same production possibility set as the aforementioned CCR and BCC models; however, they treat slacks directly in the objective function. [COOP07] The following sections provide an introduction to select non-oriented models; namely the Additive Model, Slacks Based Model and Multiplicative Model.

#### 3.6.1 ADDITIVE MODEL

Based on the BCC model, the additive model was introduced by Charnes et al. in 1985. [CHAR85a] As previously stated, the additive model is a non-oriented model that measures efficiency based on the slacks alone. The radial ($\theta$) measure of efficiency is removed and the objective function is made to maximize the sum of slacks:

**Primal Additive Model (3.23)**

Maximize: $z_0 = \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$

Subject to

$$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+$$
$$x_{i0} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-$$
$$1 = \sum_{j} \lambda_j$$
$$0 \leq \lambda_j, \text{ for } j=1,\ldots,n$$
$$0 \leq s_i^-, \text{ for } i=1,\ldots,m$$
$$0 \leq s_r^+, \text{ for } r=1,\ldots,s$$
**Dual Additive Model (3.24)**

Minimize: \[ w_0 = \sum_{i=1}^{m} v_i x_{i0} - \sum_{r=1}^{s} u_r y_{r0} + \bar{u}_0 \]

Subject to \[ \sum_{r=1}^{s} u_r y_{rf} - \sum_{i=1}^{m} v_i x_{ij} - \bar{u}_0 \leq 0, j = 1, \ldots, n \]
\[ u_r \leq \varepsilon \]
\[ v_i \leq \varepsilon \]

Under the additive model, a unit is only efficient if all slacks are zero. If a unit is deemed inefficient, its efficiency score is based on the combined input excess and output shortfall that will project the unit onto the frontier with non-negative slacks. Although the sources and amounts of a unit’s efficiency differ between the additive and BCC model, an efficient DMU under the BCC model will also be efficient under the additive model and vice versa. Additionally, the shape of the efficient frontier remains the same as the VRS BCC model. This is due to the inclusion of the convexity constraint (1 = \( \sum \lambda_j \)) in the primal model and the \( \bar{u}_0 \) variable in the dual model. Moreover, the additive model is translation invariant in both inputs and outputs meaning the translation of the original input or output values will result in the same optimal solution. [COOP07]

### 3.6.2 Slacks Based Model (SBM)

The Slacks Based Model (SBM) is an extension of the Additive model which allows the efficiency evaluation to be units invariant. This means that the model’s efficiency measure is invariant with respect to the unit of measurement of the input and output variables. Furthermore, the efficiency measure is monotone decreasing in the slacks of the input and output variables. [COOP07] The general formulation for the SBM model is as follows:

Minimize \[ \rho = \frac{1 - \frac{1}{m} \sum_{i=1}^{m} \frac{\bar{s}_i}{x_{i0}}}{1 + \frac{1}{s} \sum_{r=1}^{s} \frac{\bar{s}_r}{y_{r0}}} \] (3.25)

Subject to \[ \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- = x_{i0} \quad (i=1, 2, \ldots, m) \]
\[ \sum_{j=1}^{n} y_{rj} \lambda_j + s_r^- = y_{ro} \quad (r=1, 2, \ldots, m) \]
\[ j=1, 2, \ldots, n \]
\[ \lambda_j, s_i^-, s_r^- \geq 0 \]

Where \[ 0 \leq \rho \leq 1 \]
Under SBM, a DMU is efficient if $\rho^* = 1$ or in other words the slacks are zero. As with the additive model, the inefficient DMUs can obtain target objectives by projecting the combined input excess and output shortfall onto the frontier. This is denoted as:

$$\hat{x}_0 = x_{i0} - s_i^- (i=1,2,...,m)$$  \hspace{1cm} (3.26)

$$\hat{y}_0 = y_{r0} - s_r^+ (r=1,2,...,m)$$  \hspace{1cm} (3.27)

### 3.6.3 Multiplicative Model

The multiplicative model was developed by Charnes et al. in 1982. [CHAR82] Unlike the previously introduced DEA models, the Multiplicative Model uses multiplicative combinations for the inputs and outputs yielding a piece-wise log-linear (Cobb-Douglas type) efficient frontier. In layman’s terms, the multiplicative model is obtained by taking the anti-logs of the additive model, thus changing the summation signs ($\sum$) to product signs ($\prod$). The multiplicative model does not have a radial proportional variable and identifies inefficiencies through slack values alone. In 1983, this model was extended further, giving it non-dimensional properties. [CHAR83]

Although not extensively used, the multiplicative model offers a few opportunities to extend the range of applications for DEA. Firstly, multiplicative models do not confine efficient frontiers to concavity, but allow them to exist with both concave and non-concave regions. Furthermore, these models can provide exact estimates of elasticities. [BANK04] For a more detailed description of the multiplicative model, the reader is encouraged to consult [COOP07].

### 3.7 Results from DEA

DEA can provide several important pieces of information that can give further insight into the efficiency and performance characteristics of a DMU. The following sections discuss some key results obtained from DEA, including technical, scale and mix efficiencies, peer groups and target setting and the returns to scale of the operating environment.

#### 3.7.1 Technical and Scale Efficiency

As previously discussed, the original CCR DEA model made the assumption of constant returns to scale and excluded the effects of scale size on efficiency. The CCR model measures a unit’s overall technical efficiency which is an aggregation of technical efficiency and the effects of scale size. Therefore, any inefficiency calculated in the CCR model could be a result of inefficient technologies, non-optimal scale size or a combination of both. To separate the overall
efficiency into its components, the BCC model must be employed using the same dataset. The BCC model imposes less strict assumptions of variable returns to scale and results in a measure of technical efficiency alone. Subsequently, taking the ratio of CRS efficiency over VRS efficiency will give the unit’s scale efficiency, which measures each DMU’s distance from its optimal scale size.

Figure 3.2 provides a comparison between the CCR and BCC DEA models, with CCR represented by a dashed straight line and BCC by a solid curve. This graph can be used to ascertain information about a DMUs technical and scale efficiencies. Aside from being the only CRS efficient producer, DMU G is located on both efficient frontiers, and exhibits the highest average productivity for the given inputs. DMU G is therefore said to be operating at its most productive scale size (MPSS). This term was coined by Banker [BANK84] and refers to a DMU that is operating at the highest output per unit of input and that has the highest slope. For a more detailed look into MPSS, the reader is encouraged to review Cooper et al.. [COOP07]

Referring back to Figure 3.2, it is noted that the BCC frontier is comprised of DMUs F, G, H and I. All of these units are therefore considered technically efficient, but only those that also lie on the CCR frontier can be considered scale efficient. It follows that DMU G is the only unit that is both technically and scale efficient within this set of DMUs.

3.7.2 Mix Efficiency

Both technical and scale efficiency are radial measures and thus do not include any non-radial inefficiencies such as the slacks based measure (SBM) of efficiency. Mix inefficiencies are calculated by dividing the SBM efficiency score ($\rho$) by the CCR efficiency score ($\theta$). Mix efficiency will have a value between zero and one, with one denoting that the optimal CCR solutions have no slacks and that the model is not subject to any mix inefficiencies.

3.7.3 Peer Groups and Target Setting

DEA models also provide target setting opportunities and peer group evaluations. For each inefficient DMU, DEA provides a reference group composed of efficient peers. The target projections for each DMU will vary based on the model used, the shape of the frontier and the orientation selected. It follows that some inefficient DMUs will have multiple peers and others will only have one. The peer groups for inefficient DMUs are determined by the values given to the $\lambda$ variables. These variables represent the proportion of each efficient DMU that an inefficient DMU should be imitating.
3.7.4 Returns to Scale

As previously discussed in Sections 3.3 and 3.4, there exist both constant returns- to-scale and variable returns-to-scale DEA models. Each facet of the CCR model’s frontier has a constant slope meaning that on each facet any increase in inputs will result in a proportional increase in outputs. On the other hand, the slope of the BCC model’s frontier is allowed to change resulting in increasing, decreasing or constant returns to scale (IRS, DRS and CRS respectively). IRS occurs when marginal productivity is greater than average productivity; consequently an increase in input produces a proportionally larger increase in outputs. Conversely, DRS occur when average productivity is greater than marginal productivity.

Based on the \( \tilde{u}_0 \) variable, the BCC model can help in establishing whether a unit is operating under IRS, DRS or CRS. Moreover, under certain circumstances it is possible for the \( \tilde{u}_0 \) variable to have multiple solutions. To address this issue, Sueyoshi developed a method of identifying whether multiple solutions existed and established two unique DEA formulations. [SUEY99] This issue is also discussed in Banker and Thrall’s paper on estimating returns to scale. [BANK92]

Additionally, the convexity constraint \( (1 = \sum \lambda_j) \) in the BCC model can be altered to limit the frontier to specific returns to scale assumptions. If it is assumed that the scale of a DMU cannot be reduced but only increased, the convexity constraint becomes \( 1 \leq \sum \lambda_j \) resulting in a non-decreasing returns-to-scale model (NDRS). In an NDRS model, the frontier shows only increasing or constant returns-to-scale. Similarly, a non-increasing returns-to-scale (NIRS) model can be obtained by changing the convexity constraint to \( 1 \geq \sum \lambda_j \) resulting in only constant or decreasing returns-to-scale. A third model, known as the generalized returns-to-scale (GRS) model involves setting a lower and upper bound for the convexity constraint.

3.8 Extensions to Basic DEA Models

In order to improve model accuracy and provide a more realistic representation of the DMUs under consideration several extensions and modifications to the aforementioned DEA models are now available. From the available extensions, a select few, namely non-discretionary variables, categorical variables and restricted multiplier models, have been employed in this study to broaden the range of compatible data types and to provide a means of restricting multipliers within realistically feasible limits. These extensions are discussed in detail in the following sections.
3.8.1 Non-Discretionary Variables

In order to accurately represent a DMU’s production process, all pertinent variables affecting this process should be included in the model. However, there are cases when one or more pertinent variables are not controllable by management; thus including them in the model in the same manner as the controllable variables could produce nonsensical results. For example, the surrounding geographical environment of a bank branch is not something that management can control; nonetheless, these factors can have a great impact on the productivity, performance and services of the branch and therefore should be included in the study.

These uncontrollable variables, known as non-discretionary variables, can be dealt with in a relatively simple manner. Non-discretionary variables should be removed from the objective function of the linear program but included in the constraints. This assures that their presence is still accounted for and that their values remain constant while only the discretionary variables are optimized. For public sector analyses, Ruggiero [RUGG96] offers a modified DEA model that properly controls for non-discretionary variables that affect production by insuring that the reference groups provided by the model are feasible. This is guaranteed by constructing a reference set that includes only those DMUs with at least as harsh an environment as the one that the DMU in question faces. This is achieved by constraining the multipliers of DMUs with more favourable environments to zero.

For additional information about the inclusion and treatment of non-discretionary variables, the reader is encouraged to consult Cooper et al. [COOP07]

3.8.2 Categorical Variables

There are yet other situations when a pertinent variable has a discrete value from a set of two or more discrete values. Introduced by Banker and Morey [BANK86], these variables are referred to as categorical variables and, amongst other things, can include the presence of ATM units, weekend opening availability of branches and categorical representation of branches (i.e. rural, urban, metropolitan). There are two classifications of categorical variables, controllable and uncontrollable, which must be dealt with in different ways.

Controllable categorical variables refer to those which were selected by management. For example, management may decide to rate bank branches as poor, average and good based on their service hours. In order to fairly assess branches based on these categorical variables, they are only compared against branches that fall in the same or worse performing categories. (i.e. average branches would only be compared to average and poor branches) This is achieved by
implementing a linear programming algorithm. This algorithm also allows the DEA model to produce a unique frontier for each hierarchical level providing the ability to evaluate each DMU against different frontiers. Additionally, the recommended reference set is restricted to DMUs which exist within the same hierarchical level.

As with controllable categorical variables, non-controllable categorical variables can be handled using hierarchies. A similar linear programming algorithm can be used to ensure that inefficient DMUs are only compared to efficient DMUs operating in the same or worse environments. Again, this would produce separate frontiers and would restrict the reference set of each DMU to those within its hierarchical level.

It should be noted that these methods should not be used for non-comparable categories. Instead, either a separate analysis should be performed, or cultural adjustment factors should be used. The use of cultural factors is discussed in detail in [YANG02]. For additional insight into categorical variables refer to Cooper et al. [COOP07]

### 3.8.3 Fractional Data

A major pitfall of DEA occurs when indices, ratios and percentages are incorporated into the input or output set with other volume measures. Aside from the possible loss of information and the implicit assumption of constant returns to scale between the numerator and the denominator of the ratio, mixing of ratios and volumes presents added complexities in the scaling between the volume measures and the ratios. Moreover, the convexity constraint may not hold in such circumstances. To deal with these issues, Dyson et al. [DYS001] offer three possible solutions. First, is to use a proxy measure to replace the ratio. Second, to scale the ratio measure by a volume measure to make it compatible with the other volume measures included in your input/output set. Lastly, Dyson et al. suggest that if possible separate the numerator and the denominator and include them as an input and an output, respectively. It should be noted that each of these methods are only appropriate in certain circumstances. When using ratios, indices or percentages one should proceed with caution and carefully consider the relationships that exist between the variables in the data set.

### 3.8.4 Indices

Much like non-discretionary and categorical variables, indices must be carefully incorporated into the DEA model to ensure that issues with scaling and convexity do not arise. [DYS001] Regardless of these issues, indices are commonly used in DEA analyses to quantify
environmental, demographic or operational conditions. [PARA12] Many government agencies and organizations use indices to quantify performance, growth and other characteristics of interest making them convenient sources of information and readily available for use in DEA models. Vance [VANC00] incorporated a ‘competitive index’ into her model to take into account the local competition of neighbouring bank branches. Lozano et al. [LOZA02] incorporated environmental indices along with banking variables to assess the effect of country-specific environments on the banking industry. Paradi et al. [PARA10] incorporated two cultural indices to represent two aspects of a firm’s unique operating environment.

3.8.5 RESTRICTED MULTIPLIER MODELS AND TRADE-OFFS

DEA does not require a priori knowledge about the input and output variables but assigns multipliers to each DMU’s variables such that the DMU looks as efficient as possible. However, there are situations where the assigned multipliers do not properly represent what is realistically feasible. In these cases, the use of additional information to impose informed multiplier restrictions can result in more accurate efficiency estimates and more realistic depictions of relative efficiency and best performers. These multiplier restrictions can be based on managerial and organizational factors, physical or production limitations or any other constraint affecting the feasible range of inputs and outputs.

The concept of multiplier restriction was first introduced by Dyson et al. [DYSO88] who imposed upper and lower bounds on each multiplier. Subsequently in 1989, Charnes et al. [CHAR89] established the Cone-Ratio Method which restricted the feasible region of various multipliers to polyhedral convex cones that were defined by non-negative direction vectors. Additionally, Thompson et al. [THOM90] introduced the Assurance Region Method which imposed constraints in terms of absolute numbers or in terms of ratios assigned to the multipliers. These constraints are represented mathematically as:

**Absolute Numbers (3.28)**

\[ L_i \leq v_i \leq U_i \]

Where:  
\( L_i \) is the upper bound  
\( U_i \) is the lower bound

**Ratio Constraints (3.29)**

\[ L_{1,2} \leq \frac{v_2}{v_1} \leq U_{1,2} \]

Using the absolute number form of the Assurance Region Method is only possible when actual prices or levels of prices are known. Alternatively, the ratio constraint form can be used when the specific values are unknown but a general range of values is known.
Extending this work, Podinovsky [PODI04] suggests that the dual terms in the constraints of the envelopment model could be interpreted as production trade-offs that represent feasible simultaneous changes to the inputs and/or outputs of the technology. This trade-off approach ensures that the radial target of any inefficient unit is technologically realistic and thus the efficiency measures retain their meaning of extreme radial improvement factors. This methodology can be thought of as a technology based approach instead of the traditional value based approach.

If properly employed, restricted multiplier models can be very useful in providing more accurate and realistic modeling and can help avoid zero multipliers. However, weight restrictions must be carefully incorporated into DEA models to ensure the feasibility of results and to avoid zero or negative efficiency scores. Podinovski ascertains that the problems that arise when using weight restrictions are caused by inducing free or unlimited production of outputs in the underlying technology. In these cases the weight restrictions should be reassessed. [PODI13] For additional information and detailed descriptions of the aforementioned restricted multiplier methods the reader is encouraged to consult Cooper et al. [COOP07]

### 3.8.6 Extended Facet Efficiency Analysis

When one wishes to validate or modify the imposed cone restrictions or when multiplier restrictions and production trade-offs are not known a priori, Extended Facet Efficiency Analysis (EXFA) can be used to impose restrictions on the MPPS to ensure that the efficient frontier is made up solely of desirable facets. [OLES03]

In order to implement this methodology the facial structures of the possibility set must first be identified and then used to impose restrictions on the frontier. As detailed by Olesen and Petersen [OLSE03], the complete identification and classification of all facial structures of an empirical possibilities set can allow for the characterization of the underlying data generation process, can provide an estimation of isoquants and relative elasticities of substitution and can provide additional information on whether there is a local bias in efficiency scores. More importantly for the purpose of this study, the identification of the facial structures can be used to specify more appropriate constraints on the virtual multipliers in a Cone Ratio model and aid in the identification of outliers.

For the purpose of this study, the frontier was restricted to Fully Dimensional Efficient Facets (FDEF). FDEFs are facets which consist of exclusively strongly efficient units and have a dimension equal to s+m−1. In their recent book chapter, Olesen and Petersen detail a method in
which the DEA frontier is restricted to a FDEF in one linear program. [OLES13] The CCR and BCC forms are outlined in the linear programs below:

**CCR-EXFA (3.30)**

Maximize: \( \sum_{r=1}^{s} u_r y_{r0} \)

Subject to:
\[
\sum_{r=1}^{s} u_r y_{rf} - \sum_{i=1}^{m} v_i x_{ij} + s_j = 0 \quad j \in E
\]
\[
\sum_{i=1}^{m} v_i x_{i0} = 1
\]
\[
s_j - b_j M \leq 0
\]
\[
\sum b_j - (|E| - (s + m - 1)) \leq 0 \quad j \in E
\]
\[
b_j \text{ binary}, s_j \geq 0, u_r \geq \epsilon, v_i \leq \epsilon
\]

**BCC-EXFA (3.31)**

Maximize: \( \sum_{r=1}^{s} u_r y_{r0} + \tilde{u}_o \)

Subject to:
\[
\sum_{r=1}^{s} u_r y_{rf} - \sum_{i=1}^{m} v_i x_{ij} + \tilde{u}_o + s_j = 0 \quad j \in E_{BCC}
\]
\[
\sum_{i=1}^{m} v_i x_{i0} = 1
\]
\[
s_j - b_j M \leq 0
\]
\[
\sum b_j - (|E| - (s + m)) \leq 0 \quad j \in E_{BCC}
\]
\[
b_j \text{ binary}, s_j \geq 0, u_r \geq \epsilon, v_i \leq \epsilon
\]

Where \( E \) and \( E_{BCC} \) are the sets of strongly efficient DMUs for the given PPS.

**3.8.7 Selective Proportionality**

Prior to using DEA, an assumption about the returns to scale of the technology must be made. The CRS model requires the strong assumption of full proportionally between the inputs and outputs. It is not surprising that this proportionality cannot always be assumed even if there is a subset of outputs and inputs that are in fact proportional. In this case using a VRS model would ignore the additional information known about the variables and would effectively overestimate the efficiency scores. To provide more accurate efficiency measures in these situations Podinovski [PODI12] developed selective proportionality DEA. These hybrid DEA models allow a selected set of inputs and outputs to be treated under CRS assumptions while the remaining variables fall under VRS assumptions. These models allow more complex variable relationships to be considered and exhibit better discrimination than the VRS model.
3.8.8 **Sub-Vector DEA**

Sub-vector DEA is used to generate technical efficiency measures for a single input (output) or a subset of inputs (outputs) rather than for the entire vector of inputs (outputs). [FARE94] For example, the input oriented sub-vector efficiency of an input, $x$, is calculated by assuming that it is possible to reduce $x$ while holding the remaining inputs and outputs constant. Technical sub-vector efficiency for input variable $k$ can be calculated for each DMU $i$ by solving the following programming problem:

$$
\begin{align*}
\text{Min } & \theta^k \\ \text{s.t} & \quad -y_i + Y \lambda \geq 0 \\
& \quad x^{v-k}_i - X^{v-k} \lambda \geq 0 \\
& \quad \theta^k x^k_i - X^k \lambda \geq 0 \\
& \quad \delta x^f_i - X^f \lambda = 0 \\
& \quad N1' \lambda = 1 \\
& \quad \lambda \geq 0 \\
& \quad 0 < \delta \leq 1
\end{align*}
$$

Where $\theta^k$ is the input $k$ sub-vector technical efficiency score for firm $i$, which represents the maximum reduction of input $k$ while holding all other inputs and outputs constant. The terms $x^{v-k}_i$ and $X^{v-k}$ refer to the input vector with the $k^{th}$ column excluded.

Although there exists no sub-vector DEA bank branch analyses in literature, it has been widely applied in the agriculture sector proving its success in real world applications. Some of these applications include Oude et al. [OUDE02], who used sub-vector DEA to compare efficiency and productivity of conventional and organic farms in Finland; Oude et al. [OUDE04] who performed a sub-vector non-parametric analysis of the product of pesticide use in the Netherland; and Chebil et al. [CHEB12] who used sub-vector DEA to assess the efficiency of irrigation water use in collective irrigated schemes in Tunisia.

3.8.9 **Malmquist DEA**

In order to ensure continual improvement and growth, organizations must be aware of their productivity and how it changes over time. The Malmquist index [MALM53], not only offers a method of measuring productivity change over time but also provides information pertaining to how these changes in productivity correlate to advancements in technology and the
Bank Branch Growth Potential and Growth Trends

Malmquist indices can be used in conjunction with numerous performance measurement techniques including DEA. The most commonly used DEA based Malmquist indices are the adjacent Malmquist productivity index [FARE94a] and the base period Malmquist index [BERG92], [FORS93]. Unlike the adjacent index, the base period index offers the desirable characteristic of transitivity or circularity; however this comes at the cost of dependency on an arbitrary fixed base period. This may become problematic when technological changes occur rapidly or when long periods of time are studied. To avoid this issue but maintain circularity, Global Malmquist indices are sometimes used. ([ASMI07], [PAST05]) These global indices generate a single measure of productivity change and are not susceptible to LP infeasibility making it a desirable alternative to the more traditional methods.

Measuring productivity growth involves separating the changes in pure productivity (i.e. the level of inputs necessary to produce a level of outputs) from the changes in the relative efficiency of the DMUs over time. The Malmquist productivity index is able to measure a DMU’s productivity growth between any two periods, t₁ and t₂, by combining the measurement of the relative distances of the DMU from the frontiers, measured using the Färe distance function, and the relative changes in the position of the frontier from one period to the next. Aside from offering an overall index of productivity change, Malmquist indices can also be decomposed into a ‘catching up effect’; the change in efficiency of the individual DMU over time; and the technological change or frontier shift; the shift in the efficient frontier over time typically attributed to technological change. The mathematical formulation and decomposition for the adjacent Malmquist Index and Pastor et al.’s global Malmquist index are as follows:

Consider N production units (DMUs), n=1,...,N, that are each observed in T time periods t=1,...,T. The units are using M inputs, \( x_m \in \mathbb{R}_+^M \), to produce S outputs, \( y_s \in \mathbb{R}_+^S \). The input and output matrices, \( X^t \) and \( Y^t \) for each time period are of the dimensions \( M \times N \) and \( S \times N \), respectively. Let \( \bar{X} \) and \( \bar{Y} \) be the vectors of input matrices \( \{X^t\} \) and output matrices \( \{Y^t\} \), respectively, for t=1, ..., T. A contemporaneous production technology is then defined as \( L^t = \{(x^t, y^t)|x^t \text{ can produce } y^t\} \) where \( L^t \) satisfies free disposability, and the standard assumptions of being non-empty, closed, and convex. ([SHEP70],[FÄRE88]) Moreover, \( L^t \) satisfies constant returns to scale such that if \( x^t \in L^t(y^t) \) then \( \lambda x^t \in L^t(\lambda y^t) \) for all \( \lambda \geq 0 \).

The distance between a DMU with \((x', y')\) and the frontier at time t for an input oriented model can then be defined as:

\[
D^t(x', y') = \sup \left\{ \phi: \frac{x'}{\phi} \in L^t(y') \right\}, \quad \phi > 0
\] (3.33)
It should be noted that $D^t$ is simply the inverse of the DEA input efficiency scores. Consequently, the adjacent input oriented Malmquist index [FARE94] is defined as:

$$M^A(x^t, y^t, x^{t+i}, y^{t+i}, L^t, L^{t+i}) = \left[ \frac{D^t(x^t, y^t)}{D^t(x^{t+i}, y^{t+i})} \frac{D^{t+i}(x^t, y^t)}{D^{t+i}(x^{t+i}, y^{t+i})} \right]^{1/2}$$ (3.34)

This can be decomposed into catching up or efficiency change:

$$EC(x^t, y^t, x^{t+i}, y^{t+i}, L^t, L^{t+i}) = \frac{D^t(x^t, y^t)}{D^{t+i}(x^{t+i}, y^{t+i})}$$ (3.35)

and frontier change or technical change:

$$TC^A(x^t, y^t, x^{t+i}, y^{t+i}, L^t, L^{t+i}) = \left[ \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^t(x^t, y^t)} \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^{t+i}(x^{t+i}, y^{t+i})} \right]^{1/2}$$ (3.36)

Similarly, the global Malmquist index is defined using the distance function, $D^G$, between a DMU and the global frontier with the technology set $L^G = conv\{L^1(y) \cup ... \cup L^T(y)\}$. Where $conv\{\}$ denotes a convex set.

$$M^G(t, t + i; (\vec{x}, \vec{y})) = \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^t(x^t, y^t)} \times \left\{ \frac{D^G(x^{t+i}, y^{t+i})}{D^{t+i}(x^{t+i}, y^{t+i})} \times \frac{D^t(x^t, y^t)}{D^G(x^t, y^t)} \right\}$$ (3.37)

$$TC^G(t, t + i; (\vec{x}, \vec{y})) = \left\{ \frac{D^G(x^{t+i}, y^{t+i})}{D^G(x^t, y^t)} \times \frac{D^t(x^t, y^t)}{D^G(x^t, y^t)} \right\}$$ (3.38)

Note: The efficiency change components of both the adjacent and global Malmquist indices are calculated in the same way.

Similarly, the output oriented Malmquist models can be defined as:

$$S^t = \{(x^t, y^t) | x^t \text{ can produce } y^t\}$$

$$D^t(x^t, y^t) = \inf \left\{ \theta: (x^t, \frac{y^t}{\theta}) \in S^t(y^t) \right\}, \quad \theta > 0$$ (3.39)

Consequently, the adjacent output oriented Malmquist index [FARE94] is defined as:

$$M^A(x^t, y^t, x^{t+i}, y^{t+i}, S^t, S^{t+i}) = \left[ \frac{D^t(x^t, y^t)}{D^t(x^{t+i}, y^{t+i})} \frac{D^{t+i}(x^t, y^t)}{D^{t+i}(x^{t+i}, y^{t+i})} \right]^{1/2}$$ (3.40)

This can be decomposed into catching up or efficiency change:

$$EC(x^t, y^t, x^{t+i}, y^{t+i}, S^t, S^{t+i}) = \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^t(x^t, y^t)}$$ (3.41)

and frontier change or technical change:

$$TC^A(x^t, y^t, x^{t+i}, y^{t+i}, S^t, S^{t+i}) = \left[ \frac{D^t(x^t, y^t)}{D^t(x^{t+i}, y^{t+i})} \frac{D^{t+i}(x^t, y^t)}{D^{t+i}(x^{t+i}, y^{t+i})} \right]^{1/2}$$ (3.42)

Similarly, the global Malmquist index is defined using the distance function, $D^G$, between a DMU and the global frontier with the technology set $S^G = conv\{S^1(y) \cup ... \cup S^T(y)\}$.

$$M^G(t, t + i; (\vec{x}, \vec{y})) = \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^t(x^t, y^t)} \times \left\{ \frac{D^{t+i}(x^{t+i}, y^{t+i})}{D^G(x^{t+i}, y^{t+i})} \times \frac{D^t(x^t, y^t)}{D^G(x^t, y^t)} \right\}$$ (3.43)
\[ TC^G(t, t + i; (\bar{X}, \bar{Y})) = \left\{ \frac{D_t^{t+i}(x_t^{t+i}, y_t^{t+i})}{D_t^G(x_t, y_t)} \times \frac{D_t^G(x_t, y_t)}{D_t^{t+i}(x_t^{t+i}, y_t^{t+i})} \right\} \] (3.44)

Note: The output oriented Malmquist indices were used for the purpose of this study.

### 3.8.10 Meta Frontiers

In DEA, DMUs that do not operate under the same production technology are said to be non-comparable and thus using DEA to provide relative efficiency measures would not produce meaningful results. To ameliorate this issue Battese et al. [BATT02] developed and later improved upon [BATT04], [ODON08], a stochastic metafrontier model that was able to compare homogeneous subgroups to estimate technical efficiencies. In this technology, the group frontiers; restricted technology frontiers defined by a DMU subgroup’s operating/environmental conditions; are enveloped by a metafrontier; the boundary of an unrestricted technology set. Effectively this allows for the decomposition of the efficiencies measured relative to the metafrontier into two components; a common technical efficiency measured from the point to the group frontier and a component measured from the group frontier to the metafrontier which represents the restrictive nature of the production environment. Metafrontiers can be used in conjunction with both parametric and non-parametric methodologies. The technical efficiencies derived from this methodology provide insight into possible performance improvements which would involve changes to the management and structure of the firm. While the technology gap estimates between the group frontiers and the metafrontier can be used to determine performance improvements related to the production environment itself.

### 3.8.11 Rolling Window Analysis and Panel Models

In order to reduce dimensionality and track the changes in a DMU’s relative efficiency over time, Charnes et al. [CHAR85b] proposed a technique called “window analysis” which assesses the performance of a DMU over time by treating each DMU as a separate entity in each time period. This approach is based on defining a ‘window’ of ‘p’ periods within which each unit is treated as a separate entity and assessed by running the model for each window involving ‘np’ units. When a new time period is introduced, the earliest time period is removed creating a new window. In some cases researchers have modified this traditional methodology by removing the poorest performing period instead. This allows the new period to be compared to the best of the previous periods thereby aiding in the benchmarking process. [TALL97]

Similar to window analyses are panel models which employ panel or longitudinal data to monitor the changes in a large number of DMUs for a relatively short period of time. Panel
models provide rich environments for the development of estimation techniques and theoretical results as they incorporate both cross-sectional and time-series data. [GREE11]

3.9 **STRENGTHS AND LIMITATIONS OF DEA**

DEA offers several advantages over other commonly used performance measures, and, as with all analysis methods, has some intrinsic weaknesses. The following sections provide a synopsis of these strengths and limitations.

### 3.9.1 STRENGTHS

1. **Ability to handle multiple inputs and outputs:** Given a sufficient number of degrees of freedom; DEA models can handle any number of inputs and outputs. Other approaches, including regression analysis and SFA, can handle only multiple inputs or multiple outputs, but not both. (i.e. one input and multiple outputs or vice versa)

2. **Does not require a priori specification of functional form:** Unlike parametric analysis techniques, DEA does not require any prior knowledge of the relationships between the inputs and outputs.

3. **Does not require consistent metrics:** A DEA model is able to simultaneously handle different units without any effect.

4. **Does not require prior knowledge of variable multipliers:** The DEA model assigns variable multipliers and thus does not require them to be pre-specified. Moreover, multiplier ratios and range constraints can be applied to a model to provide more accurate and realistic results.

5. **Provide target setting capabilities:** DEA compares DMUs with a peer or combination of peers to obtain the DMU’s efficiency score. Should a DMU be inefficient, it can be projected onto the efficient frontier to provide realistic target setting.

6. **Provides a single performance score:** DEA produces a simple comprehensive efficiency score that characterizes a DMU’s relative performance.

### 3.9.2 LIMITATIONS

1. **Does not account for random error:** DEA does not account for measurement error or other noise but assumes that all deviations from the frontier are caused by inefficiencies. If random error is present, DEA may produce inaccurate results.
2. **Unable to accurately model small sample sizes:** When the degrees of freedom of a DEA model are too low (i.e. too few DMUs and too many variables), the model will tend to produce higher than normal average efficiency scores with many units appearing on the frontier.

3. **Provides a relative efficiency score:** DEA provides relative efficiency scores that are based on the specific DMUs studied. Should an important or highly efficient DMU be excluded from the analysis, the scores provided will not be as accurate.

4. **Outliers can heavily influence results:** Outliers operating with unfair advantages can greatly skew resulting efficiency scores. Other DMUs will experience lower overall efficiency scores which, in reality, are not significant.

5. **It is retrospective:** DEA is a retrospective performance measure that does not allow for future projections.
Chapter 4:
DEA and Banking

This chapter provides a detailed discussion about the use of DEA in the banking industry, including a literature survey of DEA branch analyses and an introduction to commonly used models. Commonly used extensions and relevant methodologies are also discussed.

4.1 DEA in Bank Branch Performance Evaluation

Bank branch analyses, in many cases, are more desirable and informative than institutional level analyses. Branch analyses help in understanding the complex relationships and variables that can play a role in branch level efficiency. These analyses also resolve measurement problems that exist in standard bank level analysis. [BERG97] Furthermore, the bank’s largest operational expenses are usually incurred at the branch level. Consequently, branch analyses can provide an effective managerial tool and a more direct cost management solution. Lastly, branches are the source of a large portion of the value added banking performed by customers, who, in these tough economic times, are looking for reliable face to face contact with their bank. [BCG10] Branch analysis can provide great insight into branch operations, leading to more sound and comprehensive branch strategies. If properly executed, these strategies can provide the improvement and branch network growth that is necessary to remain competitive in the financial services marketplace.

Since its conception, DEA has become one of the most widely used approaches to measure the efficiency of financial institutions. [BERG97] However, the majority of DEA banking studies have focused on banks at an institutional level, rather than at the branch level. This can partially be attributed to the difference in data availability. The majority of banks are publicly traded on major exchanges and thus, must provide their investors with quarterly and annual financial reports. This makes the collection of data for institutional level analyses rather easy. On the contrary, branch level data is mostly proprietary information and is not generally disclosed to the public. Instead, it is either amassed into bank financial reports or not reported at all. Nonetheless, surveys have shown that there has been a steady increase in DEA branch studies, nearly doubling in the last five years alone. [PARA12]

To date, there are four survey papers that review DEA applications in the banking industry, of which three focus on bank level applications and one focuses on branch level
applications. The first to review the major efficiency techniques used in the evaluation of bank performance were Berger and Humphrey. [BERG97] This survey reviewed a total of 103 papers including 57 DEA based papers. Of these, 42 focused on bank level analysis while the remaining 15 focused on branch level analysis. Berger [BERG07] provided a review and critique of over 100 studies that compared cross-national bank efficiencies obtained using various frontier techniques. Fethi and Pasiouras [FETH10] presented a review of 196 studies which employed operational research and artificial intelligence techniques to assess bank performance. Of these studies, 151 of them used DEA-like techniques and 30 focused on evaluating efficiency at a branch level. Most recently, Paradi and Zhu [PARA12] published a survey that focused heavily on the use of DEA in branch analysis. Among the 285 bank related DEA publications identified, 90 focused on branch analysis and were discussed in greater detail.

This section summarizes the key DEA model types and objectives used in branch level analyses. A brief introduction to a few common of model variations is also provided. It should be noted that the proposed study evaluates the branch network of one of the top five Canadian banks by applying certain models and extensions discussed herein.

4.1.1 MODEL OBJECTIVES

Branch level DEA applications can have a diverse set of business objectives; however, the majority of applications focus on evaluating branch specific operations. These studies allow for the exploration of efficiency determinants and provide the capability of identifying deficiencies in areas that are controllable by branch managers. That being said, branch performance measurement is not a simple task. Branches come in an assortment of sizes, operate in different economic regions and offer a variety of services to a diverse range of customers. In order for a branch performance analysis to be significant and reliable, it should capture the critical aspects of the bank’s internal operating processes, leading to a more adept understanding of these processes. Moreover, the analysis should provide target setting through the identification of best- and worst-practices and offer the capability of investigating the sources of the inefficiencies.

Depending on the objective of the analysis, different DEA model frameworks exist. Of these models, there are three that are very commonly used in branch analysis; intermediation, production, and profitability. Additionally, the market model is introduced below.
4.1.1.1 Intermediary Model

The intermediation approach was the earliest DEA model used to assess the performance of banks. [COLW92] It evaluates the branch’s ability to collect deposits and other funds (inputs) and then lend the money in various forms, including loans, mortgages and other assets (outputs). An early example of this approach is found in Alhadeff’s study (1954) that measured output in terms of dollar values of earning assets. [COLW92]

There are relatively few branch analyses that employ the intermediation approach. The first study, done by Athanassopoulos, used non-interest and interest costs as inputs and non-interest income and total volume of accounts for loans, savings and deposits for outputs. [ATHA97] Subsequently, Athanassopoulos developed a two stage model whose first stage evaluated service efficiency. The output targets produced by this stage were then used as the inputs into the second-stage intermediation model. This modified intermediary model produced a larger efficient frontier, providing each DMU with more room for improvement. [ATHA00]

4.1.1.2 Production Model

The production model, first used for branch analysis in 1985 by Sherman and Gold [SHER85], is the most popular approach for bank branch performance analysis. It views the bank branches as producers of services and products, using labour and physical resources (inputs) to produce transactions, such as loans and deposits (outputs). The transactions considered can consist of face-to-face interactions at the branch, back office transactions and delivered transactions. Consideration of customer satisfaction has also been found to play a role in the production efficiency of a branch. There have been numerous studies that have employed this methodology, some examples of which include Parkan [PARK87], Schaffnit et al. [SCHA97], Athanassopoulos [ATHA98], Camanho and Dyson [CAMA05], Portela and Thanassoulis [PORT07], Giokas [GIOK08], Tsolas [TSOL10] and Paradi et al. [PARA11].

4.1.1.3 Profitability Model

The profitability model is another commonly used method of bank branch analysis. It measures the branch’s ability to use labour, assets and capital to generate profits, or more simply put, the branch’s ability to convert expenses into revenues. The expenses (inputs) include employee expenses, occupancy expenses, branch cross charges, and other operational expenses. Additionally, loan losses and sundry can be included on the input side to penalize those branches with higher losses, whether from risky lending or unwanted revenue charges. The branch’s
revenues (outputs) are based on all of the branch’s lines of business: non-interest revenues and bank fees, interest earning from wealth management, home mortgages, consumer lending, consumer deposits, commercial lending and commercial deposits, and commission revenues earned through wealth management, credit cards and insurance brokerage. Some examples of studies employing bank branch profitability analysis are Manandhar and Tang [MANA02], Al-Tamini and Lootah [ALTA07], McEachem and Paradi [MCEA07], Paradi et al. [PARA10],[PARA11] and Tsola [TSOL10].

4.1.1.4 Market Model

Market efficiency was first defined by Athanassopouloous [ATHA95] in a study related to retail organizations. The market model measures the extent to which a bank branch, given its capacity and available resources, realizes its potential to sell products and provide services in a given market. The main objective of a bank branch is to penetrate its market by selling financial products to new customers while continuing to deliver services to existing customers. To achieve market efficiency, a bank branch must expand their outputs and optimize how cohesively the branch (size, number of employees, etc.) fits into its market conditions. Examples of market models applied to bank branch efficiency analysis include Athanassopouloous [ATHA98], Thanassoulis [THAN99], Athanassopouloous and Giokas [ATHA00], and LaPlante [LAPL15].

4.1.2 Model Variations

DEA offers a wide range of model variations that must be carefully considered when building a model for a particular bank branch analysis. Inevitably, modellers must determine which specific DEA version to apply and which extensions are necessary to properly evaluate the given data and meet the analysis objectives. In this section, several model variations and relevant extensions are introduced.

4.1.2.1 Returns to Scale

In bank branch efficiency analysis, DEA can be employed assuming either constant returns to scale (CRS) or variable returns to scale (VRS). The first DEA model introduced by Charnes et al. [CHAR78] was a CRS model referred to as the CCR model. This model has been used extensively in research, accounting for nearly half of all bank branch efficiency studies. [PARA12] McEachern and Paradi [MCEA07] state that when commercial and specialty branches, such as oil and gas, and real estate branches, were excluded, branches operate on a CRS basis. Fethi [FETH10] on the other hand suggests that CRS models are only suitable when
all DMUs are operating at an optimal scale. Approximately 30% of branch analyses employ the BCC model, a VRS approach. Cook et al. [COOK00] used VRS models to examine the efficiency of over 1300 Canadian branches that have multiple resource sharing operating functions. Camanho and Dyson [CAMA05] claimed that the efficient frontier should be estimated assuming VRS for the production model and CRS for the value-added approach. The remaining 20 or so percent of branch studies employed both CRS and VRS methods in their analysis. Ultimately, determining the most suitable model form is dependent on both the data set and the analysis objectives.

4.1.2.2 INPUT ORIENTED VS. OUTPUT ORIENTED

DEA can estimate efficiency under either an input-oriented or output-oriented approach. The input-oriented approach determines by how much input quantities can be reduced without changing output quantities. Conversely, the output-oriented approach determines by how much output quantities can be increased without changing input quantities. Branch managers tend to have more control over the inputs (labour and capital) than the outputs (profit, loans and transactions) and thus, input oriented approaches are more commonly used in branch performance efficiency studies. It should be noted that there also exists a more complex non-oriented approach called the slack-based measure that looks to simultaneously decrease inputs and increase outputs. [FETH10]

4.1.2.3 MULTI-STAGE DEA ANALYSIS

In the business world, it is not uncommon to witness managers employing unilateral strategies that are purely in response to the head office’s assessment criteria. This improvement approach tends to lead to decreased efficiency in areas not accounted for in the strategy. Bank branches are complex, multi-faceted organizations that require careful assessment and personalized strategies. In attempts to provide a more comprehensive performance measure, several researchers have simultaneously applied more than one DEA model to evaluate overall branch efficiency. Moreover, this methodology provides better target setting opportunities. Examples of multi-level DEA models used for branch analysis include Sherman and Ladino [SHER95], Manandhar and Tang [MANA02], Jahanshahloo et al. [JAHA04], Al-Tamimi and Lootah [ALTA07], Portela and Thanassoulis [PORT07], Giokas [GIOK08] and Paradi et al. [PARA11].
4.1.2.4 Relevant Extensions

When performing a branch efficiency analysis using DEA, there are several model extensions that are required in order to produce realistic and attainable targets. Of these extensions, restricted multiplier models, non-discretionary variables and fractional data are introduced herein. The theoretical properties of these extensions are discussed in more detail in Chapter 3.

Restricted Multiplier Model

As previously mentioned, DEA does not require a priori knowledge of the functional form or of the variables used. Although this is a desirable characteristic, it can potentially result in solutions with variable multipliers that are not realistically feasible. In these situations, it can be very useful to employ multiplier restrictions. This extension proves very valuable from a managerial standpoint as it provides managers with the opportunity to limit the changes made to any specific variable. Examples of restricted multiplier models applied to bank branch analyses include Schaffnit et al. [SCHA97], Cook and Hababou [COOK01], Paradi and Schaffnit [PARA04] and Paradi et al. [PARA11].

Non-Discretionary Variables

There are numerous instances when a variable is highly pertinent to the evaluation of a DMU but cannot be controlled by management. These variables, referred to as non-discretionary variables, can include anything from weather to population demographics to branch square footage. The very nature of these variables prevents their alteration or manipulation, and thus, they cannot be treated in the same manner as discretionary variables. To avoid unachievable recommendations from the DEA model, non-discretionary variables are generally removed from the objective function in the linear programming model, but are included in the constraints so that their presence is still accounted for when evaluating the DMUs. Non-discretionary variables are very common in all types of DEA efficiency analyses including bank branch analysis. Some more recent studies employing non-discretionary variables in branch analysis include Paradi and Schaffnit [PARA04], Wu et al. [WU06b], Portela and Thanassoulis [PORT07] and Paradi et al. [PARA11].
Fractional Data

When evaluating banks and bank branches, there are many cases where the most readily available data are in the form of ratios. In order to use this data to create a successful DEA model, one must keep in mind a few key points. Firstly, the use of ratios leads to the loss of information about the size of a unit, and implicitly assumes constant returns to scale in the operation of the DMUs in question. Many of these performance indicators use different denominators, and thus, have the virtue of being independent of size. The ratio approach will not lead to major difficulties, provided that the process of assigning multipliers is closely monitored. They should continue to represent the value of a unit increase in one ratio relative to a unit increase in another. However, if the ratio form of DEA is used, then it is crucial that the BCC form be implemented instead of the CCR form. [HOLL03]

Rolling Window Analysis and Panel Models

Selection of an appropriate performance measurement technique can be highly dependent on the type and availability of data. Improvements in data collection capabilities and increased transparency in banks has led to the availability of substantial data sets that span over large periods of time. Although traditional techniques can provide insight into performance at specific cross sections of time, they are not able to quantify the dynamic changes that occur within the branch network. It is here that Window Analyses and Panel Models are useful in examining the effectiveness of both variable heterogeneity and time on economic processes.

If data are available over consecutive years, it can be valuable to employ window analysis and panel models in conjunction with DEA techniques to improve dimensionality and to monitor the variations in the efficiency scores of the DMUs over time. Since its introduction by Charnes et al. [CHAR85b], window analysis has been employed in numerous DEA analyses. However, due to the proprietary nature of bank branch data, bank branch DEA window analyses are much less common. Some of the more recent DEA window analysis studies related to banking efficiency include Webb’s study which utilized window analysis to investigate the relative efficiency levels of large UK retail banks during the transition period of 1982-1995 [WEBB03], Asmild et al., who demonstrated that when Malmquist indices are based on DEA window analysis scores, inconsistent results may be obtained [ASMI04], and Fadzlan et al. who employs DEA window analysis to determine the relationship between Singapore banking efficiency and stock returns [SUFI07]. Panel models face similar restrictions due to data availability but can be
modified depending on the completeness of the data. Some examples of DEA panel models used for bank branch analysis include Tahir et al. [TAHI09] and Staub et al. [STAU09].
Chapter 5:  
Banking Data: Background, Collection and Treatment

Bank branch analyses are often more desirable than institutional analysis since they can provide management with tailored branch level target objectives and cost management strategies and more informative efficiency and profitability measures. However, branch analyses require extensive amounts of detailed proprietary branch data which is generally difficult to access and is, more often than not, incomplete when finally obtained. Moreover, researchers rarely have a choice in what data they are provided directly from the bank and thus must rely on several sources of data to achieve a complete data set. Post collection, careful data cleaning and treatment must be performed to ensure the validity of the analysis results. Both data collection and treatment are onerous, yet crucial tasks that play integral roles in performing a determinative bank branch performance analysis. As such, a large amount of work went into collecting and treating data prior to the formulation and application of either of the developed methodologies and played an equally important role in the development of both methodologies. Hence, this section first introduces all data sources, including the Canadian and Turkish banks' that provided data for this analysis. All data collection and treatment techniques are then be discussed and the development of all indices included in this analysis are summarized. The R scripts associated with this chapter are provided in Appendix B.

5.1 CANADIAN AND TURKISH BANKING SYSTEMS

The legislations and regulations that govern the financial and banking institutions of a country have a large impact on the operations, services and the growth potential of those institutions. In order to objectively examine the performance of these institutions, one must fully understand the laws that govern them. This section introduces both the Canadian and Turkish banking systems and provides a brief overview of the banks that are assessed in this study.

5.1.1 CANADIAN BANKING SYSTEM

The banking system in Canada is regulated by the Canadian Bank Act (S.C. 1991, c.46), last revised in 1991 by the Office of the Superintendent of Financial Institutions (OSFI). [DJC14] Established in 1987 under the ‘Office of the Superintendent of Financial Institutions Act’, OSFI is a federal agency that regulates and supervises federal financial institutions and private pension
plans. OSFI is responsible for assuring that institutions are working in sound financial conditions and are complying with their governing laws. Additionally, OSFI contributes to the development and administration of regulations and legislations. [OSFI13]

Currently, OSFI regulates 77 banks; 23 are domestic, 26 are foreign bank subsidiaries and 23 are full-service foreign banks. Banks in Canada are further categorized into Schedule I, II and III. Schedule II and III respectively deal with foreign bank subsidiaries and the bank branches of foreign institutions. Schedule I encompasses all of the domestic banks. [CBA14]

Canadian Banks, as well as certain Schedule II banks, are eligible for deposit insurance provided by the Canadian Deposit Insurance Corporation (CDIC), a federal crown corporation. Created by Parliament in 1967, the CDIC insures deposits up to $100,000 per depositor in each member institution and reimburses depositors for the amount of any insured deposits should any member fail. [CDIC12] The CDIC helps to maintain the stability of the Canadian financial system.

As of April 30th, 2014, Canada’s domestic banks held approximately $3.8 trillion with the top five banks, referred to as the “Big Five”, accounting for nearly 93% of the assets held by Canadian deposit-taking institutions. [OSFI13] The Big Five, which include BMO Financial Group (BMO), Bank of Nova Scotia (BNS), Canadian Imperial Bank of Commerce (CIBC), RBC Financial Group (RBC) and TD Canada Trust (TD), operate in what some call an oligopoly. Small domestic and foreign banks and new entrants are as yet little competition for these market dominant institutions; however, there is substantial competition amongst themselves. The Big Five’s shares are widely held, with any entity allowed to hold a maximum of twenty percent. [DJC14] In the second quarter of 2014, the “Big Five” alone held $2.2 trillion in total loans, $2.3 trillion in deposits, and $203 billion in total shareholders’ equity. [CBA14]

Canadian banks service a wide range of customers including individuals, small and medium-sized businesses, large corporations, and governments. They offer a variety of banking, investment and financial services through large branch networks. Canada is considered to be one of the most heavily “branched” countries in the industrialized world, with more than 6,150 full service branches. Moreover, Canadian banks employ in excess of 270,750 people and provide the service of over 18,500 Automated Banking Machines (ABMs). [CBA14]

5.1.2 “CANADIAN BANK” OVERVIEW

One of the “Big Five” Canadian banks, the Canadian Bank evaluated in this study ranks within the top 75 banks worldwide in terms of asset size. [CBA14] The Bank includes a branch
network of more than 1,000 branches and employs over 40,000 branch and corporate personnel. The Bank practices retail, commercial and corporate banking and offers an extensive range of financial products and services. The following table (Table 4.1) provides a partial list of products and services offered by the Bank. These services are administered through several channels including in-branch, debit cards, Automated Banking Machines (ABMs), and internet and telephone banking.

Table 5.1: Canadian Bank’s Retail and Commercial Products and Services

<table>
<thead>
<tr>
<th>Retail and Commercial Products and Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Bank Accounts (Chequing, Savings)</td>
</tr>
<tr>
<td>- Lines of Credit</td>
</tr>
<tr>
<td>- Mortgages and Other Loans</td>
</tr>
<tr>
<td>- Mutual Funds and Investment Services</td>
</tr>
<tr>
<td>- Investment Banking and Brokerage</td>
</tr>
<tr>
<td>- Credit Card</td>
</tr>
<tr>
<td>- Foreign Exchange, Wires and Bank Drafts</td>
</tr>
<tr>
<td>- Insurance Brokerage</td>
</tr>
</tbody>
</table>

5.1.3 TURKISH BANKING SYSTEM

Over the past three decades, the Turkish banking system has undergone substantial changes that have greatly impacted both the banking and the financial sectors of Turkey. Prior to the 1980s, the Turkish economy was structured as a planned economy where the state had a majority share in almost all areas of economic activity. In order to strengthen the banking system, substantial changes were made to the Turkish Banking Law in 1985. These changes introduced new requirements pertaining to capital and problem loans, and improved accounting, reporting and deposit insurance standards. Additionally, external auditing became mandatory for banks and regulatory barriers restricting entry of new banks into the banking system were lifted. In 1987, the Central Bank of Turkey began open market operations, and a year later the foreign exchange market was established and interest rate controls were abolished. Foreign exchange trading and capital movements were fully freed in 1989. [KESK00]

Since the reform in the 1980s, the Turkish Banking sector has shown great improvements, however it is still in a stage of development. The limited growth experienced by the sector can largely be attributed to the low level of savings due to low income levels, prolonged high inflation rates and the low demand for financial assets due to high intermediation costs caused by heavy taxation. These tax burdens have also lead to the predominance of government securities in Turkey’s capital markets. [BAOT05]

The banking system in Turkey accounts for roughly 90 percent of the total assets of institutions in the financial sector. Although banking institutions dominate the financial sector, there has been a recent increase in the number of non-bank financial institutions which include...
special finance institutions, insurance companies, leasing companies, consumer finance companies, intermediary institutions in the capital market, real estate investment trusts, and private pension funds. These institutions are governed by several supervisory authorities including the Capital Markets Board of Turkey, the under secretariat of Treasury under the Prime Ministry of the Republic of Turkey, and the Banking Regulation and Supervision Agency.

Banks in Turkey can be classified under one of two distinct groups; commercial banks, which have the permission to collect deposits and non-depository banks, which, as the name implies, do not accept deposits. These groups can be further sub-divided into state-owned, privately owned, and foreign banks depending on source of their capital.

Commercial banks offer a wider range of products and services including traditional depository and lending services, investment banking and capital market transactions. There are currently 33 commercial banks, 3 of which are state owned, 13 are privately owned and 16 are foreign banks. The state owned commercial banks have wide networks and specialize in financing the agricultural sector and SMEs. The privately owned banks are large-scale commercial banks that provide various services and that engage more in wholesale banking. Lastly, the foreign banks are classified as either those who have opened a branch in Turkey or those who have been founded in Turkey. Turkey’s non-deposit banks perform mainly capital market transactions, portfolio management and consulting. There are currently 13 non-deposit banks; 3 that hold public capital, 8 with private capital and 2 with foreign capital. [BAOT13]

To ensure the stability of the banking sector and insure client deposits, the Saving Deposit Insurance Fund (SDIF) functions to collect the receivables of failed banks. If the financial structure of a bank becomes weakened it is taken over by the SDIF along with all of their liabilities.

Much like Canadian banks, Turkish banks service a wide range of customers including individuals, small and medium-sized businesses, larger corporations, and governments. They have a vast branch network of over 10,000 branches that employ approximately 185,000 employees and hold over 1.2 trillion dollars' worth of assets. [BAOT13]

5.1.4 “TURKISH BANK” OVERVIEW

The Turkish bank examined in this study is one of the top ten Turkish banks. It offers a branch network of approximately 750 branches and employs over 13,000 branch and corporate personnel. The Bank is a commercial bank and thus offers a wide range of products and services including traditional depository and lending services and investment banking. The following
Table (Table 4.2) provides a partial list of products and services offered by the Bank. These services are administered through several channels including in-branch, debit cards, Automated Banking Machines (ABMs), and internet and telephone banking.

Table 5.2: Turkish Bank’s Retail and Commercial Products and Services

<table>
<thead>
<tr>
<th>Retail and Commercial Products and Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Bank Accounts (Chequing, Savings)</td>
</tr>
<tr>
<td>- Lines of Credit</td>
</tr>
<tr>
<td>- Mortgages and Other Loans</td>
</tr>
<tr>
<td>- Mutual Funds and Investment Services</td>
</tr>
<tr>
<td>- Investment Banking and Brokerage</td>
</tr>
<tr>
<td>- Credit Card</td>
</tr>
<tr>
<td>- Foreign Exchange, Wires and Bank Drafts</td>
</tr>
</tbody>
</table>

5.2 DATA COLLECTION

As proprietary bank branch data is often times very difficult to obtain, the direction of a branch study is often driven and limited by the data that is available. Consequently, researchers must first assess their data sets and identify what variables paint a sufficiently complete picture of what they wish to study. Often, additional data must be collected from external sources so that a comprehensive and cohesive representation of branch operations can be achieved. For this study, two data sets were obtained from banking institutions, one Canadian and one Turkish. Each of these data sets included exclusively product and fund related data thus lending themselves well to use in a productivity model framework. In addition, data was collected from Statistics Canada and other publically available sources to develop environmental and competitive indices for inclusion in the Canadian data set.

This section details each data set as well as any additional data collection that was performed. The creation of a simulated data set for use with the restricted growth model is also discussed. It should be noted that all data sets were used in conjunction with the restricted model, while only the Canadian data set contained enough time periods and variables for use in the growth model. To ensure clarity, this section is separated by country.

5.2.1 CANADIAN DATA

Banking Data

The Canadian bank provided a substantial data set, with upwards of 70 variables, including categorical (region, market etc.) and non-categorical (products, assets etc.) variables as well as postal codes and transit numbers for over 1000 of its branches. Employment, which was
not provided by the bank, was collected from the Pitney Bowes Canadian Business Data file [PITN10] and matched to the banking data via transit number.

It should be noted that within the data set there are cases of branch overlap which occur due to the fact that all of a customer’s transactions, regardless of where they were carried out, are linked to the original branch with which they opened their accounts. Unless detailed customer transaction data are provided, there is no way to account for this within the proposed models. However, the existence of branch overlap should not affect the results as the purpose of this study is to measure the growth of a branch from one year to the next. In most circumstances, perhaps with large metropolitan areas being an exception, customers base their choice of branch on the location most convenient to them or some other desirable quality such as customer service or an attractive incentive scheme. Regardless of why they have made their choice, these models give credit to branches for gaining more customers from one year to the next. The reasons behind this growth can later be identified through individual branch analysis if the Bank so desires.

*Competing Branch Data*

Through the aggregation of the Bank’s and Pitney Bowes’ data sets, GPS (Global Positioning System) coordinates for each of the Bank’s branches were also provided. Moreover, the Pitney Bowes file included GPS locations for all bank branches for every bank in Canada allowing the number of competing branches in a given area to be calculated. This was done for a 50km, 25km, 10km, 5km and 1km radii, resulting in unique measures of local competition for each of the Bank’s 1000 plus branches. These values were employed as the competitive index in this study.

*Statistics Canada Data*

In order to ensure a realistic and unbiased performance measure, local environmental factors that affect a branch’s ability to perform should be included in the analysis. A branch’s ability to grow is heavily reliant on the size of its potential customer base as well as its local economy. To account for these factors, 2006 and 2011 population and income data was collected from Statistics Canada, accessed through the University of Toronto’s CHASS system, for all (approximately 65,000) Dissemination Areas (DAs) across Canada. [STAT13] (GPS coordinates were also collected for each DA.) This data set was carefully cleaned and the values for both income and population were projected based on DA or region specific growth trends for the years 2008, 2009, 2010, 2012 and 2013. The GPS coordinates of each branch were then used to
determine the 50 closes DAs from which weighted population and income indices were created for each branch.

GDP was not chosen as an environmental variable for several reasons. First and foremost, although GDP is often seen as an indicator of quality of life, it is not always a good proxy for absolute level of standard of living as it can increase with a decrease in average household income. Often GDP does not account for the subtle economic and environmental fluctuations that may impact consumer banking behaviours whereas population and average income are directly related to the region’s observed spending and saving patterns. [COST09] Moreover, GDP for each DA was not available and thus unique GDP indices could not be formulated for each branch. The use of income and population indices provides more detailed information about the operating environments of each branch and allows one to study the impact of both variables on a branch’s ability to grow.

5.2.2 **Turkish Data**

A second data set was acquired from a large Turkish bank which included staffing, product and customer data for over 750 branches measured at monthly intervals for the year of 2012. Although extensive, no information pertaining to branch location was provided within this data set and thus environmental and competitive indices could not be created for these branches.

5.3 **Data Treatment**

In order to construct the most meaningful models and produce valid results, the data must be critically analyzed and the variables for each model should be carefully selected. This was accomplished through the application of various techniques including correlation and sensitivity analyses, and the removal of outliers by various techniques. Segmentation for use in local analyses was also performed. Refer to Appendix A for a step-by-step description of the data treatment and outlier removal performed in for this analysis.

5.3.1 **Variable Selection and Sensitivity Analysis**

When selecting variables for inclusion in a specific model, it is important to keep the model’s objective in mind as well as avoid redundancy in the variables. Reducing the number of highly correlated variables (proxies) helps increase the discriminatory power of DEA without adversely affecting the final efficiency scores. Moreover, model misspecification caused by the
omission of relevant variables and inclusion of irrelevant variables can be detrimental to a DEA analysis. [SMIT97] In this study, correlation analysis was used to aid with variable selection.

**Correlation Analysis**

Correlation analysis is a very commonly used technique for DEA variable selection. It employs the use of the correlation coefficient ($\rho$) which measures how two variables vary together. The coefficient’s value can range from +1, denoting a perfect positive linear relationship, to -1, denoting a perfect negative linear relationship, with 0 representing no correlation at all. The correlation coefficient can be calculated using the following formula:

$$\rho_{v_1,v_2} = \frac{\text{cov}(v_1,v_2)}{\sigma_{v_1}\sigma_{v_2}} = \frac{E[(v_1-\mu_{v_1})(v_2-\mu_{v_2})]}{\sigma_{v_1}\sigma_{v_2}}$$  \hspace{1cm} (5.1)

Where:
- $v_1,v_2$: are the variables being compared
- $\text{cov}(v_1,v_2)$: is the covariance of $v_1$ and $v_2$
- $\sigma$: represents the standard deviation of the specified variable
- $E$: is the expected value operator
- $\mu$: the mean of the specified variable

Correlation analysis, in terms of variable selection, requires that the correlation coefficient be calculated for every combination of inputs and outputs. When there are two inputs (or outputs) that are highly correlated with one another ($|\rho|>0.95$), this suggests that the variables are too similar and convey essentially the same information. It follows that one of the two should be removed, or the two should be combined together. On the other hand, when comparing an input variable with an output variable, there should be some degree of correlation ($>0.50$) between the two, ensuring that they are both statistically significant. The results obtained from correlation analysis should be used with caution as high correlation does not mean with certainty that a variable should be removed. Additionally, there are instances were leaving a variable in the model makes more sense managerially.

**Efficiency Contribution Method**

Sensitivity analyses can provide further insight into which variables highly influence the results of a DEA model and which do not. For the purpose of this study, two forms of sensitivity analysis were used. First, the efficiency contribution method was used to determine the influence of each variable on the average DEA scores. Each variable was individually removed and the model was re-run with the remaining variables. The difference in mean scores between the re-run
model and the original model were then compared to determine whether the variable has a significant impact on the DEA scores achieved.

**Principal Component Analysis**

Principal Component Analysis (PCA) employs a similar methodology to the efficiency contribution method, but is a more widely accepted technique for variable selection. PCA employs the use of the Wilcoxon Rank-Sum test which is a non-parametric statistical hypothesis test that is used when comparing two related samples to assess whether their population mean ranks differ. It can be used when the population cannot be assumed to be normally distributed making it an ideal way of determining whether the efficiency distribution of a DEA model post-variable removal is equivalent to that of the original model. [COOP07]

Cooper et al. [COOP07] offer a simplified version of the test where the two independent data sets of efficiency scores are defined as \( A = \{a_1, a_2, ..., a_m\} \) and \( B = \{b_1, b_2, ..., b_n\} \). The data from A and B are combined to form a set of \( m+n \) observations in a new data set referred to as D. This new data set is then ordered from largest to smallest and ranked, with any identical values in D receiving a mid-rank (sum of ranks/number of identical values). The sum (S) of the ranks of either data set A or B is taken and S is then normalised using the following equation:

\[
T = \frac{(S - m(m+n+1)/2)}{\sqrt{(mn(m+n+1)/12)}} \tag{5.2}
\]

According to Cooper et al., \( T \) has an approximately standard normal distribution. Using the value obtained for \( T \), the null hypothesis that the two groups have the same population at a level of significance of \( \alpha \) can be checked. The hypothesis is rejected if either \( T \leq -T_{\alpha/2} \) or \( T \geq T_{\alpha/2} \).

Although this methodology is relatively simple when dealing with small data sets, the requirement of mid-ranks becomes increasingly difficult and time consuming for large datasets. For the purpose of this study, the R function ‘wilcox.test’, which employs the traditional Wilcoxon Rank-Sum Test, was used. Like its simplified version, the null hypothesis of the Wilcoxon Rank-Sum Test is that the two groups have the same population. A significance level of \( \alpha = 0.05 \) was used for this test and thus the hypothesis was rejected if either \( Z \leq -1.96 \) or \( Z \geq 1.96 \). The \( \rho \)-statistic, also provided by the R function, is a non-parametric measure of overlap between the two data sets and is widely used in studies of categorization. It can range in value from 0 to 1 with both extreme values representing the complete separation of the distributions and 0.5 representing complete overlap.
5.3.2 **Outlier Removal**

The results of a DEA model can be heavily skewed by the existence of outliers, which can include DMUs with erroneous data, and non-comparable DMUs (i.e. oil and gas branches, and real estate branches). To ensure the credibility of the model's results, it is essential that the data be cleaned of outliers. Several methods of outlier removal are used in conjunction with DEA including stripping the efficient frontier, statistical data analysis, manual cleaning and Wilson’s outlier detection statistic. [WILS03] These methods, each of which were used to some degree within this work, are summarized herein.

**Stripping the Efficient Frontier**

DEA provides relative efficiency measures and peer references that are based on the efficient units that compose the efficient frontier. It follows that the outliers that find themselves on the frontier can affect the efficiencies of numerous units and thus pose the greatest risk to the soundness of the results. Stripping the Efficient frontier is a quick but crude way to determine if the DMUs that lay on the frontier are in fact outliers. To do this all of the DMUs on the efficient frontier are removed and the model is rerun. The difference in the average efficiency scores and the number of efficient units are then assessed. This method provides a quick way to determine if there are any obvious outliers on the frontier. However, for more precise results other methods have also been used. Given the availability of branch specific information pertaining to branch type and location, outlier removal should be performed on an individual branch basis rather than by means of frontier stripping.

**Statistical Data Analysis and Manual Cleaning**

Statistical data analysis is another simple method of outlier identification which entails calculating the coefficient of variation for each variable used within the model by dividing the variable’s standard deviation by its mean. The coefficient of variation is a normalized measure of the dispersion of a variable with respect to the mean of its population; higher coefficients of variation indicate that there exists greater variation within the variable’s values. The histogram of the variable with the largest coefficient of variation is then examined to determine which DMUs have extreme values. These DMUs are subsequently removed from the analysis.

The most straightforward approach to outlier removal requires that the data be manually examined and any visible anomalies be removed. The obvious downfall to this technique is that it is a very time consuming process and thus not always feasible. It does, however, allow for
more exact outlier removal. The data used in this study was carefully examined and any units with noticeable data errors or inconsistencies were removed.

**Wilson’s Outlier Detection Statistic**

The Wilson Outlier Detection Statistic is a statistical methodology used for identifying outliers by means of ranking DMUs in terms of their dissimilarity to other DMUs. It is often applied to production sets with multiple inputs and outputs that are used in deterministic non-parametric frontier models.[WILS03] Several highly regarded benchmarking software packages include this methodology as their main method of outlier detection. ([BOGE13],[WILS13])

Theoretically this methodology is thought to be useful when data sets are large and manual data checking is too onerous a task, however substantial computing power is required to successfully solve the statistic for large data sets.

### 5.3.3 Segmentation

Whether performed pre- or post-analysis, the segmentation of DMUs based on individual or multiple characteristics can provide new and useful insights about the DMUs in question. For the purpose of this study, DMUs were segmented based on demographic characteristics prior to the application of the period growth model. To determine the ideal number of clusters, the widely accepted heuristic approach involving a visual inspection of the within group sum of squares vs. the number of clusters graph was used. As can be seen in Figure 5.1 the sum of the squares decreases monotonically as the number of clusters increases. However, from some number of clusters onwards (between 6 or 7 in Figure 5.1) the decrease flattens markedly creating an “elbow” which indicates the appropriate number of clusters to be used. [MILL85]

![Figure 5.1: Within Group Sum of Squares vs. Number of Clusters](image)
The segmentations were then determined using k-means segmentation; one of the most readily used cluster analysis techniques. [WU08] This technique segments DMUs by grouping then with ‘like’ DMUs by partitioning n observations (n DMUs) into k clusters in which each observation belongs to the cluster with the nearest mean. This methodology can be used for single- or multi-variable segmentations.

The significance and usefulness of the results obtained from a DEA analysis are highly dependent on whether or not the DMUs are, in fact, comparable. Although the DMUs used for this study are under the same regulations and corporate culture, there may be other factors, such as local environment, that impact the operations of the branch. Therefore, to better assess the efficiency of the DMUs, and to provide more realistic peer references and target setting, local analyses should also be performed. Local analyses offer the opportunity to compare units only to those that are most comparable. The DMUs are segmented based on a chosen characteristic or variable and then local DEA analyses are run for each segment, where only the DMUs in that segment are included. This generally results in an overall increase in DEA efficiency scores which can be attributed to the decrease in differentiation between the DMUs within the production set.
Chapter 6: Restricted Growth Model

Generating growth from one year to the next is one of the key performance objectives of most bank branch managers. DEA bank branch analyses found in the existing literature use contemporaneous data and thus cannot address this objective. For example, several DMUs from the data sets introduced in Chapter 5 that were experiencing losses in customers, funds and products were identified to be efficient when using a traditional DEA production model. DMUs which experience negative growth should not be considered efficient as these production losses are detrimental to the Bank’s efforts to increase market share. Moreover, DEA projects the inefficient unit onto the efficient frontier and takes a convex combination of the closest efficient units in order to calculate their efficiency scores. Inefficient units then reference the misspecified efficient units resulting in skewed efficiency scores and misspecification of inefficient units as well. Consequently, disregarding growth with the use of a traditional DEA model may ultimately produce ineffectual results when it comes to achieving branch growth.

The traditional restricted models introduced in Section 3.8.5 can, in some cases, provide a partial solution to this problem; however, they cannot definitively ensure that negative growth units will not appear on the efficient frontier. Super-efficiency models, which are used to provide better discrimination between efficient units [LOVE03], offer the concept of removing efficient DMUs from the PPS when calculating their efficiency but do not provide the ability to remove select units from the PPS when calculating the efficiencies of inefficient units. It follows that there does not exist a model that can appropriately handle negative growth units that fall within the MPPS.

This chapter introduces a new restricted DEA model that imposes a constraint on the non-negative intensity variable to ensure that negative growth units cannot be referenced by inefficient units. The model objective and a comprehensive explanation of the methodology are provided, along with three detailed applications of the new technology whose results are discussed. All R scripts associated with this chapter are provided in Appendix C, Section 1.

6.1 Objective

The objective of this model is to provide a means of restricting negative growth DMUs from being in the reference technology set of a DEA model. This new technology ensures that
the bank’s objective of maintaining positive growth is properly considered within the performance analysis of its branch network by restricting negative growth units from falling within the MPPS. Moreover, it corrects the efficiency scores of inefficient units by assigning them more appropriate peer groups consisting exclusively of positive growth units. In summation, the model proposed herein provides an efficiency analysis that is able to account for growth unlike the traditional productivity model.

6.2 METHODOLOGY

As stated in the objective, the goal of this model is to restrict a specific set of units from being either in the MPPS or from being reference units for inefficient units. To achieve this, a restriction must be imposed on the non-negative intensity variable of the DMUs in question within the dual or envelopment model of DEA. The input oriented VRS model is provided in equation 6.1, where \( G_j \) represents growth. (Note that the exposition is identical for the CRS model given the convexity constraint is removed.)

Minimize: \[ \theta \] 
Subject to: 
\[ \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r0} \quad r = 1, \ldots, s \quad u \in \mathbb{R}^s_+ \]
\[ \theta x_{i0} - \sum_{j=1}^{n} x_{ij} \lambda_j \geq 0 \quad i = 1, \ldots, m \quad v \in \mathbb{R}^m_+ \]
\[ -\sum_{j=1}^{n} \lambda_j = -1 \]
\[ \lambda_j G_j \geq 0 \quad j = 1, \ldots, n \quad \kappa_j \]
\[ \lambda_j \geq 0, j = 1, \ldots, n \]

For the purpose of this study, a value of 0 or 1 was assigned to \( G_j \) for each DMU depending on whether the unit experienced negative or non-negative growth, respectively. A unit was determined to have negative growth if there was a loss in customer and/or funds from one year to the next. This strict definition was chosen to ensure that both customer and fund growth were taken into consideration as requested by the bank. Depending on the application of this methodology, the definition of \( G_j \) can be modified to represent any characteristic that one would like to base their restriction on or a threshold that must be surpassed. As shown in equation 6.1, if growth is non-negative, the DEA model progresses as usual and the non-zero constraint on the intensity variable remains unchanged. If the growth for a particular unit is zero, however, the intensity variable for that unit is given a value of zero. This means that this unit cannot be
referenced by any other unit in the PPS. Moreover, it means that the unit itself must reference other units to determine its own efficiency score.

If one wishes to calculate the multipliers for each input and output the primal DEA formulation must be implemented. The procedure for the input oriented primal VRS model is detailed in equation 6.2. Note that this equation uses the index set $G^{pos} = \{ l | G_l \geq 0 \}$ which consists of non-negative growth DMUs.

$$\text{Maximize: } \sum_{r=1}^{s} u_r y_{r0}$$

$$\text{Subject to: } \sum_{r=1}^{s} u_r y_{rj} - \sum_{l=1}^{m} v_l x_{lj} - \bar{u}_0 \leq 0 \quad j = 1, ..., n \quad j \in G^{pos}$$

$$\sum_{l=1}^{m} v_l x_{l0} = 1$$

$$u_r \geq 0, r = 1, ..., s$$

$$\bar{u}_0 \in \mathbb{R}$$

$$v_l \geq 0, i = 1, ..., m$$

To aid in the understanding of the consequences of these restrictions on the MPPS a simple two dimensional representation of an input oriented efficient frontier is provided in Figure 6.1.

![Graphical Representation - Restricted Growth Model](image)

**Figure 6.1: Graphical Representation - Restricted Growth Model**

Let unit B be a DMU with negative growth that, in a traditional DEA model, finds itself on the efficient frontier. For clarity, these units are referred to as ENG (Efficient Negative Growth) units from here on in. Much like a super efficiency model, the restricted model removes the unit from the MPPS consequently moving the frontier from ABC to AC. The model then
calculates unit B’s efficiency by projecting it onto the new efficient frontier; effectively calculating the super-efficiency for the unit. This efficiency measure should be disregarded as it does not provide a meaningful performance measure for the unit. Although this process may seem counter intuitive to the stated objective, this model does allow peer groups to be determined for ENG units. For example, unit B, despite its efforts, may not be aware of what is preventing its ability to achieve positive growth. This model provides a peer group consisting of its most similar efficient peers, units A and C, that unit B can reference for operational advice in the future. From a managerial standpoint, this may prove to be invaluable information that is not otherwise provided by a traditional DEA model.

One could interpret the treatment of growth in this restricted model as being akin to an environmental variable, where dominance is only allowed by branches with non-negative growth rates. In this sense, this methodology is similar to Ruggiero’s [RUGG96] method of dealing with hierarchical environments, but differs in the fact that branches with negative growth are not allowed to dominate any branch, even others with negative growth. This tighter constraint results in conditions where it is infeasible to solve for ENG units. In the CRS model, infeasibility occurs when, in an input-oriented model, the ENG unit has the only zero value for any input or the only positive value for any output among all of the non-ENG units in the reference set. Given the nature of the data used in DEA bank branch analyses, it is very unlikely that this type of infeasibility would occur. Infeasibility cannot arise in the output oriented CRS restricted model.

For the VRS condition, there are more possible conditions for infeasibility due to the additional convexity constraint. The pattern of zeros mentioned above for the CRS case causes infeasibility in the VRS case. Moreover, if there is no reference DMU for the ENG infeasibility will occur. These two cases can occur in either orientation. For clarity, Table 6.1 outlines orientation specific conditions for infeasibility.
### Table 6.1: Restricted Growth Model - Conditions for Infeasibility

<table>
<thead>
<tr>
<th>Model</th>
<th>Input Oriented</th>
<th>Output Oriented</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRS</td>
<td>Negative Growth Unit the only zero value for any input or the only positive value for any output among all of the ENG units in the reference set</td>
<td>N/A</td>
</tr>
<tr>
<td>VRS</td>
<td>Case 1: at least one output is strictly larger than the corresponding output for any other non-ENG unit in the reference set  &lt;br&gt;Case 2: at least one output is strictly larger than any convex combination of that output among all non-ENG units in the reference set</td>
<td>Case 1: at least one input is strictly smaller than the corresponding input for any other non-ENG unit in the reference set  &lt;br&gt;Case 2: it has at least one input is strictly smaller than any convex combination of that input among all non-ENG units in the reference set</td>
</tr>
</tbody>
</table>

Figure 6.2 provides a visual representation of the feasible range of an output oriented VRS model with two inputs. In this model, \( \theta y_{r0} \leq \sum_{j=1}^{n} y_{rj} \lambda_j \) constraint always holds for all \( \lambda_j \) and thus the model is infeasible if and only if there does not exist a \( \lambda_j (j\neq0) \) with \( 1 = \sum \lambda_j \) such that \( x_{i0} \geq \sum_{j=1}^{n} x_{ij} \lambda_j \) holds. For additional information, refer to Seiford and Zhu [SEIF99] who provide an in depth analysis of the infeasibilities associated with the super efficiency model.

![Infeasibility Graph - Output Oriented Model](image)

**Figure 6.2: Infeasibility Graph - Output Oriented Model**

Real examples of infeasibilities are provided in the application sections of this chapter. It should be noted that the presence of infeasible units does not affect the results of this model nor does it take away from its scientific contribution. As is shown in the following section, few cases
of infeasibility occur when applied to real data sets and the results obtained for the remaining units provide invaluable information and direction for improvement.

The effect of this model on inefficient units is demonstrated through unit D in Figure 6.1. In the original DEA model, the efficiency of unit D is measured by \( \text{OD'/OD} \), while the efficiency measure using the restricted model is \( \text{OD''/OD} \). It can be seen that the use of the restricted model increases the efficiency score of the inefficient unit by removing the influence of the ENG unit. The model is able to adjust the efficiency score of D to reflect its true rank within the set of DMUs. Moreover, the model assigns unit D a new peer group consisting of units A and C as opposed to units A and B. The efficiency scores obtained from the restricted model for inefficient units are either the same or greater than those obtained from a traditional DEA model. Moreover, the efficiency scores of inefficient units are only altered if their peer groups originally contained an ENG unit. It should also be noted that inefficient units may become efficient once negative growth units are removed from the MPPS. Infeasibility for inefficient units will only occur if they themselves become an ENG unit.

To test the proposed technique, an R-script was written for the primal and dual input and output oriented forms of the CRS and VRS models. (Refer to Appendix C, Section 1) The script uses a “Growth” variable to determine whether the DMU is within the set of ENG units or not. This variable was manually assigned to each DMU based on their overall growth. A value of 0 was given to those units with negative growth and a value of 1 to those with zero or positive growth.

6.3 **Demonstrative Application**

In order to provide a clear example of how the restricted model behaves when applied to a data set, a small data set including 39 DMUs was generated to simulate real banking data. To create the data set, the ranges and averages for each variable were obtained from the Canadian Banking data set. These values were used to simulate 39 real DMUs with normally distributed variables. To investigate infeasibilities, a small number of units were modified to create extreme units with either large value outputs for the given inputs, or small value inputs for the given outputs. The values for the Growth variable were then randomly assigned to each of the DMUs. This section introduces the applied model formulation used and provides a detailed review of the results and a discussion of their interpretation.
6.3.1 MODEL

In order to test the validity and applicability of the restricted model, it was applied to a traditional production model; one of the most commonly used DEA formulations. The formulation for this model is provided in Table 6.2. Combining the restricted model methodology with the production model provides information pertaining to branch production efficiency, while accounting for a branch’s growth resulting in a more comprehensive measure of branch efficiency.

Table 6.2: Test Set Model Formulation

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>Number of Customers</td>
</tr>
<tr>
<td></td>
<td>Number of Products</td>
</tr>
<tr>
<td></td>
<td>Amount of Funds Held</td>
</tr>
</tbody>
</table>

6.3.2 RESULTS AND DISCUSSION

Using the developed R-script, the restricted model methodology was applied to the test set for both the CRS and VRS formulations. The results are summarized in Table 6.3. For the complete results refer to Appendix C, Section 2.

Table 6.3: Test Set Results Summary

<table>
<thead>
<tr>
<th></th>
<th>CRS</th>
<th>Restricted CRS</th>
<th>VRS</th>
<th>Restricted VRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Efficiency</td>
<td>0.558</td>
<td>0.590</td>
<td>0.749</td>
<td>0.754</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
<td>5</td>
<td>3</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td># of ENG units</td>
<td>3</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td># of Units with</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Efficiency &gt; 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of Infeasible DMUs</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>4</td>
</tr>
</tbody>
</table>

From the results, it is apparent that the restricted model successfully performs the desired task of removing ENG units and appropriately readjusting the efficiency scores of the remaining units. It can be seen that the average efficiencies increased for both the CRS and VRS models. On further inspection, it is apparent that the restricted model in both cases has reassigned previously inefficient units as efficient. In the CRS case, 5 units were originally efficient with 3 of them being ENG units. After their removal from the MPPS, the restricted model determines another unit is operating on the efficient frontier. Similar results are observed in the VRS case.

For the CRS case, the restricted model was able to feasibly solve for all three ENG units, effectively providing peer reference groups for each. The VRS model, as could be predicted, had
a higher occurrence of infeasibility. To get a better sense of which types of units were susceptible Table 6.4 details each of the infeasible units.

**Table 6.4: Test Set Results- Infeasible Units**

<table>
<thead>
<tr>
<th>DMU</th>
<th>6</th>
<th>9</th>
<th>29</th>
<th>35</th>
<th>PPS Average</th>
<th>Min in PPS</th>
<th>Max in PPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inputs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Employee</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Output</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td># Customers</td>
<td>521</td>
<td>1770</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td># of Products</td>
<td>2605</td>
<td>8000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Amount of Funds</td>
<td>19688590</td>
<td>56478640</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Result Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRS</td>
</tr>
<tr>
<td>RCRS</td>
</tr>
<tr>
<td>VRS</td>
</tr>
<tr>
<td>RVRS</td>
</tr>
<tr>
<td>RTS</td>
</tr>
<tr>
<td>Outlier?</td>
</tr>
</tbody>
</table>

In order to determine whether these units were extreme units or outliers, the Wilson Outlier Statistic was used. The results of the analysis, obtained through the use of the FEAR package in R [FEAR13], are provided in Figure 6.3 and Table 6.5. A copy of the R-script used to run this analysis is provided in Appendix B.

![Figure 6.3: Wilson Outlier Statistic Graph](image-url)
Table 6.5: Wilson Outlier Statistics

<table>
<thead>
<tr>
<th>DMU #</th>
<th>R₀</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.527995</td>
</tr>
<tr>
<td>2</td>
<td>0.224971</td>
</tr>
<tr>
<td>3</td>
<td>0.149915</td>
</tr>
<tr>
<td>4</td>
<td>0.10214</td>
</tr>
<tr>
<td>5</td>
<td>0.066704</td>
</tr>
<tr>
<td>6</td>
<td>0.032897</td>
</tr>
<tr>
<td>7</td>
<td>0.023178</td>
</tr>
<tr>
<td>8</td>
<td>0.016138</td>
</tr>
<tr>
<td>9</td>
<td>0.010884</td>
</tr>
<tr>
<td>10</td>
<td>0.006953</td>
</tr>
<tr>
<td>11</td>
<td>0.004599</td>
</tr>
<tr>
<td>12</td>
<td>0.002831</td>
</tr>
<tr>
<td>13</td>
<td>0.001639</td>
</tr>
<tr>
<td>14</td>
<td>0.000892</td>
</tr>
</tbody>
</table>

In Figure 6.3, the large peak at 2 purports that units 19 and 35 are outliers, while the small peak at 5 suggests that units 4, 33 and 34 are possible outliers. The differences in R₀ values observed beyond point 5 are negligible and therefore further information pertaining to outliers cannot be obtained through this method. Referring to Table 6.5, it is apparent that unit 35 is an outlier based on its R₀ value resulting in its infeasibility in the restricted model. Units 9 and 29 were not clearly identified to be outliers using the Wilson statistic.

Unit 6 provides an interesting case: it is not originally efficient in either the CRS or VRS model, but becomes efficient in the restricted VRS case. However, unit 6 is a negative growth unit and therefore becomes part of the ENG set. Given that it does not have any close enough peers, the solution for unit 6 becomes infeasible. This highlights another positive attribute to this model: instead of removing negative growth units completely from the PPS, this model allows them to be considered as long as they are not influencing other units. This allows for useful target objects to be provided for all units including those with negative growth.

Through this application, it is clear that the restricted growth model is able to carry out its desired purpose on small data sets with like units. The efficiency scores of inefficient units that previously referenced ENG units were properly corrected resulting in a slight increase in average efficiency scores and new peer groups were assigned. Of the 39 DMUs investigated in this application, only 4 were found to be infeasible and all fell within the hypothesized conditions for infeasibility defined previously. Moreover, peer groups were successfully provided for all ENG units with feasible solutions.
6.4 Application to Canadian Banking Data

Although the simulated data set provided promising results, true validation of this methodology requires its application to real-world data. The Canadian Banking data set provided an ideal opportunity to test the developed methodologies on a vast and diverse set of branch data. Not only does this offer insight into the restricted model’s ability to provide realistic and usable results but it also tested the technique’s ability to deal with large data sets consisting of a wide range of DMUs of varying sizes. Moreover, it provided the opportunity to evaluate the effect of environmental variables on the model’s ability to find feasible solutions for each DMU.

This section outlines the model formulation used in conjunction with the Canadian bank's data set and discusses the results and findings obtained through this application.

6.4.1 Model

In addition to the reasoning mentioned above, data availability played a large role in model selection for the Canadian banking data. The dataset, which consists predominantly of financial product data, is best suited for use in a production model. In addition to a general production model, a production model which included environmental indices was also executed to determine whether the inclusion of environmental variables affects the validity of the results. The model formulations for both executions are summarized in Table 6.6.

Table 6.6: Canadian Banking Data Model Formulation

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>Number of Customers</td>
</tr>
<tr>
<td>Income Index*</td>
<td>Number of Products</td>
</tr>
<tr>
<td>Population Index*</td>
<td>Amount of Funds Held</td>
</tr>
<tr>
<td></td>
<td>Competitive Index*</td>
</tr>
</tbody>
</table>

* This model was run both with and without indices

6.4.2 Results and Discussion

Like the test set, the restricted model was applied to the Canadian banking data set for both the CRS and VRS formulations. The growth for each DMU was determined based on whether the DMU experienced a loss of funds or a loss of clients. This methodology was chosen based on the conviction that banks do not want to lose funds, customers or the customers’ SOW as these are key factors in maintaining financial health and market share. Of the more than 1,000 DMUs examined in this analysis 207 experienced one or both forms of negative growth. The results of the model without and with indices are summarized in Tables 6.7 and 6.8, respectively.
Due to the size of the data set, the complete results from these analyses are not provided in this document but are available on request.

### Table 6.7: Canadian Banking Data Results Summary - Without Indices

<table>
<thead>
<tr>
<th></th>
<th>CRS</th>
<th>Restricted CRS</th>
<th>VRS</th>
<th>Restricted VRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Efficiency</td>
<td>0.214</td>
<td>0.214</td>
<td>0.231</td>
<td>0.233</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td># of ENG units</td>
<td>0</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td># of Units with</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Efficiency &gt;1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of Infeasible DMUs</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 6.8: Canadian Banking Data Results Summary - With Indices

<table>
<thead>
<tr>
<th></th>
<th>CRS</th>
<th>Restricted CRS</th>
<th>VRS</th>
<th>Restricted VRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Efficiency</td>
<td>0.252</td>
<td>0.254</td>
<td>0.400</td>
<td>0.412</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
<td>21</td>
<td>18</td>
<td>43</td>
<td>41</td>
</tr>
<tr>
<td># of ENG units</td>
<td>4</td>
<td>-</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td># of Units with</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Efficiency &gt;1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of Infeasible DMUs</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>8</td>
</tr>
</tbody>
</table>

The results of this application support the findings of Section 6.3. Although there were a smaller percentage of ENG units within the Canadian dataset, they were all successfully removed from the frontier and the efficiency score properly adjusted. As shown in Table 6.7, only a small number of infeasibilities occurred for the production model without indices. Given the large size of the dataset, it is more likely that ENG units have enough close peers for there to be a solution that meets the convexity requirement.

Through the comparison of the models with and without indices, it is apparent that the addition of indices increased the probability of achieving infeasibility. First, it appears that there were 2 units that, like unit 6 discussed in the previous section, were not efficient but became ENG units when using a restricted VRS model. It is likely that there is a cluster of negative growth units that sits on the edge of the frontier. Thus when the restricted model is applied, they all become ENG units with no close peers. This cluster of units results from the use of environmental indices which cause DMUs with similar environments to be sitting close to one another in the PPS. In the case of these DMUs their local environment may be playing a large role in their inability to achieve growth. To validate this hypothesis, the units that became infeasible exclusively in the restricted VRS model were investigated. Of these 5 units, 4 were located in a cluster on the IRS portion of the frontier. It follows, that once the restricted model was applied none of these units had any close positive growth peers. This highlights the model’s ability to identify demographic regions that are more likely to house negative growth branches.
which is, from the Bank’s perspective, a very valuable insight. Unfortunately, the Wilson Outlier Statistic is highly inefficient for very large datasets and could not be executed for this model.

The application of the restricted model to the Canadian banking dataset further validated its usefulness and proved that the model has real world applicability. It is, however, apparent that indices should be used with caution in conjunction with the restricted model. When applying the restricted model to large and diverse datasets, it may be more useful to cluster the DMUs based on their environmental factors and running the without indices model on each cluster.

6.5 **APPLICATION TO TURKISH BANKING DATA**

To further investigate the consistency and stability of the restricted model, it was also applied to a data set provided by a Turkish banking institution. Not only did this provide an additional opportunity to test the validity of the methodology, but it also provided the opportunity to test the model on data from branches that have vastly different company culture and environmental conditions than those of the Canadian institution.

This section outlines the model used in conjunction with the Turkish banking data and provides a detailed discussion of the results obtained from the restricted model analysis.

6.5.1 **MODEL**

Much like the Canadian banking data set, the Turkish data consisted predominately of product data resulting in the use of a production model. This data set provided more specific product information including consumer and mortgage loans as well as demand deposits and term deposits. Unfortunately, due to the lack of location specific branch information, environmental indices could not be created for this data set. The model formulation is summarized in Table 6.9.

<table>
<thead>
<tr>
<th>Table 6.9: Turkish Banking Data Model Formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
</tr>
<tr>
<td>Number of Employees</td>
</tr>
<tr>
<td>Consumer Loans</td>
</tr>
<tr>
<td>Mortgage Loans</td>
</tr>
<tr>
<td>Demand Deposits</td>
</tr>
<tr>
<td>Term Deposits</td>
</tr>
</tbody>
</table>

A Comprehensive Study of Bank Branch Growth Potential and Growth Trends
6.5.2 RESULTS AND DISCUSSION

Following the same methodology as used in the above applications, the restricted model was applied to the Turkish data for both the CRS and VRS formulations. The growth for each DMU was determined based on whether or not the DMU experienced an overall monetary loss across all products. Of the 775 DMUs examined in this analysis, 311 experienced an overall loss. The results of this application of the restricted model are summarized in Table 6.10. Again, due to the size of the data set, the complete results from this analysis are not provided in this document but are available on request.

<table>
<thead>
<tr>
<th>Table 6.10: Turkish Banking Data Results Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Efficiency</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
</tr>
<tr>
<td># of ENG units</td>
</tr>
<tr>
<td># of Units with Efficiency &gt;1</td>
</tr>
<tr>
<td># of Infeasible DMUs</td>
</tr>
</tbody>
</table>

In terms of the average efficiency scores, the results from this application are very similar to those achieved by the previous two analyses; as are the results of the CRS model. The VRS model, however, produced some interesting results unlike those previously obtained. First and foremost it should be noted that the majority of the DMUs are operating on the IRS portion of the frontier and a very large portion of these IRS units are operating efficiently. This results in an efficient frontier with sections that are very densely packed with DMUs. Many of these efficient units were found to have negative growth (187). This presents a very interesting application of the restricted growth model. From Table 6.10, it can be seen that of the 187 ENG units identified, only 1 of them produces an efficiency score greater than 1 when the restricted VRS model is applied and only 8 are found to be infeasible. The remaining 178 ENG units maintain their efficiency score of 1, and thus remain on the efficient frontier. This anomaly occurs along the aforementioned dense patches on the frontier. Since each unit sits so closely to several peers, the LP is able to find a convex combination of peers for the ENG units that results in an efficiency score of 1. Much like the >1 efficiency scores, these should be disregarded in terms of the ranking of units as the ENG units are underperformers. However, this does provide the important insight that these ENG units have very closely related peers which are operating under nearly identical conditions but are able to produce positive growth. The information gathered from these
peers is not only credible but highly relevant, making it even more attractive from a managerial standpoint.

This application has highlighted the usefulness of the restricted model in instances where there are a large number of negative growth units. Furthermore, it has demonstrated that when the DMUs within the PPS are very similar, infeasibilities are reduced and the appropriateness and amount of information gleaned from the results can increase quite substantially.

6.6 **COMPARISON OF RESTRICTED MODEL RESULTS**

Upon comparing the results from the Canadian and Turkish applications it is apparent that there are inherent dissimilarities between each country’s branches which cause widely differing restricted model results. These dissimilarities stem from the differences that exist between the company cultures, operating environments and banking sectors of the two countries. Canada has a highly regulated banking sector that operates in a relatively stable economic environment. This results in branches that operate nearly identically under CRS. It comes as no shock then, that there are few Canadian ENG units identified. On the other hand, the Turkish banking sector is still in a stage of development and is less heavily regulated than the Canadian banking sector. It follows that the economic environment is less stable resulting in numerous IRS branches whose main focus is on perceived performance regardless of whether it is sustainable or beneficial to the Bank in the long term. It is these IRS units that make up the majority of the ENG units identified by the restricted model.

6.6 **EXFA APPLICATION AND DISCUSSION**

To further validate the usefulness and applicability of the restricted model, the EXFA model was applied to each data set and the results were compared to those discussed above. Since multiplier restrictions were not known for the data sets, the EXFA model provides an opportunity to derive realistic multiplier restrictions that are then used to restrict the original DEA model. This section details the application of the methodology and discusses how the obtained results further validate the existence of the proposed restricted model.

6.6.1 **METHODOLOGICAL APPLICATION**

To execute the EXFA methodology introduced in Section 3.8.6, an R-script was written and tested on several textbook examples prior to its application to the Test, Canadian and Turkish data sets. The completed EXFA R-script and its complementary ‘Slacks’ function script
are provided in Appendix C, Section 1. This methodology was applied to each data set using the same production models outlined above. The results of each of the EXFA applications were then compared to the results of the restricted model to determine its ability to remove ENG units from the frontier.

6.6.2 RESULTS AND DISCUSSION

Canadian Data

Due to the dimensional requirements needed to achieve a fully dimensional facet (i.e. must consist of exclusively strongly efficient units and have a dimension greater than s+m-1), the CRS model could not be executed for the Canadian data set as it has fewer than the required amount of strongly efficient units. The results of the “with” and “without” indices VRS models are provided in Table 6.11.

<table>
<thead>
<tr>
<th>Table 6.11: EXFA Canadian Banking Data Results Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRS DEA</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Without Indices</td>
</tr>
<tr>
<td>Average Efficiency</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
</tr>
<tr>
<td># of ENG units</td>
</tr>
<tr>
<td>With Indices</td>
</tr>
<tr>
<td>Average Efficiency</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
</tr>
<tr>
<td># of ENG units</td>
</tr>
</tbody>
</table>

Turkish Banking Data

<table>
<thead>
<tr>
<th>Table 6.12: EXFA Turkish Banking Data Results Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEA</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>CRS</td>
</tr>
<tr>
<td>Average Efficiency</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
</tr>
<tr>
<td># of ENG units</td>
</tr>
<tr>
<td>VRS</td>
</tr>
<tr>
<td>Average Efficiency</td>
</tr>
<tr>
<td># of Efficient DMUs</td>
</tr>
<tr>
<td># of ENG units</td>
</tr>
</tbody>
</table>

By comparing the number of ENG units identified using the traditional DEA models, the restricted model and the EXFA model, it is apparent that the restricted model successfully removes these units while the EXFA model does not. Although useful in its own right, the fully dimensional facet restriction of the EXFA model does not take into account the additional growth information and therefore cannot properly account for the yearly growth of each branch.
These results further validate the need for a DEA model that is capable of properly accounting for growth.

6.7 CONCLUSION

In order to remain competitive, banking institutions must ensure that their branch networks are operating efficiently while continually growing their customer base, their SOW and the amount of funds they hold. To aid in this effort, banks have been known to employ traditional DEA models to determine the relative efficiency of their branch networks. However, it was found that these traditional models do not properly account for negative growth branches, resulting in misspecification of units and downward skewed efficiency scores. To ameliorate this problem and provide a more comprehensive and realistic performance evaluation technique, a new restricted growth model was developed. This model imposes an additional constraint on the non-negative intensity variable given the DMU has experienced negative growth. All DMUs to which this constraint is applied cannot fall within the MPPS and therefore cannot be referenced by inefficient units. This ensures that all peer groups consist of positive growth units and all operational target objectives are associated with producing positive growth.

In order to verify and validate the model, both the primal and dual formulations of the input and output oriented CRS and VRS models were applied to three distinct data sets; one simulated, one provided by a Canadian Bank and one provided by a Turkish Bank. The results of the simulated data set were used to determine sources of infeasibility which are outlined in detail above. Although there are conditions that result in infeasible solutions, the integrity of the model and the usefulness of the results are not diminished by these infeasibilities. Inefficient units, as well as most ENG units, had provided more realistic peer groups and the efficiency scores of inefficient units are corrected to properly account for growth. Application to two distinct and extensive real-world data sets further illustrated the usefulness and broad applicability of this model as well as demonstrated the infrequency of infeasibilities in larger datasets consisting of truly comparable DMUs. These applications also highlighted the model's ability to identify demographic regions that are more susceptible to having negative growth branches. This information is invaluable for the Bank to improve its growth strategies in underperforming regions. The restricted model proposed in this study was found to consistently produce significant results and successfully met its objective of bridging the gap between traditional DEA models and proper growth accounting. Additionally, this model can easily be modified for use in other applications where one wishes to restrict efficiencies based on a specific external variable.
Despite the existence of other DEA restriction methods, such as the cone ratio or the EXFA models, these methodologies were not capable of achieving the objective of the restricted model. The cone ratio model requires a priori information about the applied restrictions limiting its applicability, while the EXFA was found inadequate at removing ENG units. It follows that the restricted growth model provides a significant theoretical contribution to the field of DEA.
Chapter 7: 
Growth Efficiency Model

Like any business or corporation, the Bank continually strives to improve growth margins and increase market share in order to remain competitive. In the current economic times, the retention and growth of customers and funds has become increasingly difficult and is a major focus of banking institutions. In order to provide a more comprehensive means of assessing growth performance, a new DEA formulation tailored to measuring the relative growth efficiency of bank branches is developed herein. This chapter outlines the motivation and objectives for this model and provides a detailed explanation of the methodology used in its development. The feasibility of assessing growth trends with the proposed methodology is then examined through the application of Malmquist technologies and rolling window analysis. For verification, a real-world application is provided, which includes a thorough examination and discussion of results. To conclude, the developed model is combined with the aforementioned restricted growth methodology for a final comprehensive growth performance measure.

7.1 MOTIVATION AND OBJECTIVE

Within recent years, banking institutions have faced significant uncertainties due to the volatile nature of financial markets. Consequently, it has become more important than ever that their branch networks are operating efficiently and are, at worst, maintaining the bank’s overall market share. Although several useful bank branch performance measurement techniques exist to evaluate productivity, profitability and intermediation, none of these measures properly assess a branch’s ability to grow its customer base, share of wallet (SOW) or funds. From a managerial stand point, it is more sensible to use a high growth branch as an example for poorly performing branches than one with high production efficiency but overall customer and asset losses from year to year. To bridge the gap between growth analysis and performance efficiency measures, a new DEA methodology was developed. This methodology incorporates production and resource data from consecutive years to provide a growth efficiency score for each branch. From this, valuable insights into how well each branch is using its resources and assets to grow its customer base and assets in the subsequent year are obtained. To date, no performance measure exist which allows the assessment of such growth directly from the DEA model.
This study then extends these methodologies to include the ability to assess growth trends for each branch. Through the application of Malmquist indices and their careful interpretation, information pertaining to branch growth trends and technology changes can be obtained. This information is vital in understanding a bank’s growth trend as well as the causality behind changes in branch growth efficiencies. In summation, the objective of this study is to provide a measure of branch growth efficiency through the development of a new DEA methodology as well as a means of assessing growth efficiency trends over an extended period of time.

7.2 METHODOLOGY

After the model objective was clearly defined, it was apparent that the new DEA model would need to properly account for the dynamic nature of growth which could only be achieved through the use of non-contemporaneous data. It follows that the development and examination of this model was highly reliant on the availability of a sufficiently extensive panel data set. Advantageously, in its completion, the Canadian data set contained customer, product and resource data for several time periods. This offered the unique opportunity of incorporating a dynamic element of change through the inclusion of two consecutive year’s data. It should be noted that time lagged inputs do appear in DEA literature. For example, Charnes et al. [CHAR95] used lagged marketing inputs in their study of brand efficiency in the carbonated beverage industry. Özpeynirci and Köksalan [ÖZPE07] developed a methodology to properly account for time lags in a DEA model. However, these lags appear solely as inputs and do not appear as outputs at any point. Moreover, the purpose of including lags in this literature is not to measure growth but to more accurately model the production process and the lags that exist between obtaining inputs and producing outputs.

Although the proposed growth efficiency methodology can be applied to any of the traditional DEA models, the available data lent itself more appropriately to the production model. The formulation for the growth efficiency model is provided in Table 7.1.

**Table 7.1: Growth Model**

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources (t+1)</td>
<td>Resources(t)</td>
</tr>
<tr>
<td>- FTEs</td>
<td>- FTEs</td>
</tr>
<tr>
<td>Products(t)</td>
<td>Products(t+1)</td>
</tr>
<tr>
<td>- Amount of Funds</td>
<td>- Amount of Funds</td>
</tr>
<tr>
<td>- Number of Products</td>
<td>- Number of Products</td>
</tr>
<tr>
<td>- Number of Customers</td>
<td>- Number of Customers</td>
</tr>
<tr>
<td>- Etc.</td>
<td>- Etc.</td>
</tr>
</tbody>
</table>
Traditional DEA models employ contemporaneous data and thus measure productivity efficiencies at a specific time. These models do not factor in the time lapse that exists between obtaining resources and using them for production operations, nor do they consider the production numbers from the previous year as being pertinent to their future production efficiency. Although the efficiency scores obtained from traditional models do provide insight into the overall operational efficiency of the DMUs, these models do not consider any dynamic aspects of production. To ameliorate these issues, the proposed model incorporates a dynamic element of change through the inclusion of product data at time ‘t+1’ as an output and at time ‘t’ as an input. To complete the model, resource data from time ‘t+1’ is included as an input. This method of incorporating data from consecutive time periods effectively allows for the calculation of a branch’s growth efficiency from period ‘t’ to ‘t+1’. (i.e. its ability to grow the number of customers and the amount of funds and products over a period from ‘t’ to ‘t+1’ given the number of customers, funds and products that existed in ‘t’.) Branches whose production in period ‘t’ is high and whose production in period ‘t+1’ is low will perform worse than those with low period ‘t’ production and high production in period ‘t+1’.

This methodology provides a framework for estimating growth efficiency. Applying this framework requires one to make an assumption about the returns to scale. For the application presented herein both CRS and VRS assumptions have been used, as bank branches are commonly argued to operate under both of these assumption in literature. The mathematical formulations for the primal output-oriented CCR and BCC models are provided below:

**Primal CCR Output-Oriented Model (7.1)**

Minimize:  
\[
\sum_{i=1}^{m_x} p_i x_{i0,t+1} + \sum_{k=1}^{m_y} p_k y_{k0,t} 
\]

Subject to 
\[
\left[ \sum_{r=1}^{s_y} q_r y_{r,t+1} + \sum_{t=1}^{s_x} q_t x_{t,t} \right] - \left[ \sum_{i=1}^{m_x} p_i x_{ij,t+1} + \sum_{k=1}^{m_y} p_k y_{kj,t} \right] \leq 0
\]
\[
\sum_{r=1}^{s_y} q_r y_{r0,t+1} + \sum_{t=1}^{s_x} q_t x_{t0,t} = 1 
\]
\[
-p_i^x \leq -\varepsilon; -p_k^y \leq -\varepsilon; -q_t^x \leq -\varepsilon; -q_r^y \leq -\varepsilon 
\]
Primal BCC Output-Oriented Model (7.2)

Minimize: \[ \sum_{i=1}^{m^x} p^x_i x_{i0,t+1} + \sum_{k=1}^{m^y} p^y_k y_{k0,t} - \bar{\nu}_0 \]

Subject to

\[ \left[ \sum_{r=1}^{s^y} q^y_r y_{r,t+1} + \sum_{\ell=1}^{s^x} q^x_{\ell} x_{\ell,t} \right] - \left[ \sum_{i=1}^{m^x} p^x_i x_{ij,t+1} + \sum_{k=1}^{m^y} p^y_k y_{kj,t} \right] - \bar{\nu}_0 \leq 0 \]

\[ \sum_{r=1}^{s^y} q^y_r y_{r0,t+1} + \sum_{\ell=1}^{s^x} q^x_{\ell} x_{0\ell,t} = 1 \]

\[ -p^x_i \leq -\varepsilon; -p^y_k \leq -\varepsilon; -q^x_{\ell} \leq -\varepsilon; -q^y_r \leq -\varepsilon \]

\[ \bar{\nu}_0: \text{free in sign} \]

Where:

- \( p^x \) are the input weights given to the inputs at time ‘t+1’
- \( p^y \) are the input weights given to the outputs at time ‘t’
- \( q^x \) are the output weights given to the inputs at time ‘t’
- \( q^y \) are the output weights given to the outputs at time ‘t+1’
- \( m^x \) is the number of inputs at time ‘t+1’
- \( m^y \) is the number of outputs at time ‘t’
- \( s^x \) is the number of inputs at time ‘t’
- \( s^y \) is the number of outputs at time ‘t+1’

As depicted in Figure 7.1, this methodology can be extended over a longer time period for a more comprehensive evaluation of branch growth.

![Figure 7.1: Consecutive Year Period Growth Analysis](image)

In order to fully appreciate the results of this model, the cause of the variation that exists between the growth efficiencies of branches must be investigated. This may include market conditions, local environment and demographics, the branch's reputation in the community, or the ability to access new clientele.

The proposed methodology provides a straightforward method of calculating growth efficiency which allows for the use of existing DEA software. For the purpose of this study, the highly regarded Benchmarking [BOGE13] and FEAR [WILS13] R-packages were used.
7.3 **Growth Trends**

In addition to measuring growth efficiency, it is very desirable that the proposed methodology work in conjunction with trend analysis techniques. Trend analysis can provide additional information linking growth efficiencies with sources of causation thus making it a crucial component in executing a comprehensive branch growth analysis. Since each DEA analysis deals with a unique PPS, results from two distinct analyses cannot be directly compared and thus different trend analysis techniques must be tested and their results verified. For the purpose of this study, the developed formulation was used with the two most commonly used DEA trend analysis techniques; namely Malmquist indices and rolling window analysis. This section provides a brief synopsis of the application of these techniques and the interpretation of their results.

### 7.3.1 Malmquist Indices

The Malmquist index, introduced in Section 3.8.8, is the classic means of measuring TFP growth when using DEA. It provides the ability to assess the changes in productivity through the use of relative distance functions. This methodology also offers the ability to decompose the index into a frontier shift component and an efficiency catch up component. The frontier shift component measures the technological change that results in a shift of the frontier from one period to the next while the catch up component assesses the change in a DMU’s position relative to the frontier from one period to the next. The focus of this study is to use these decompositions in order to assess both the global and local changes in growth and provide a measure of relative growth performance for each individual DMU. To test the feasibility of this methodology both adjacent and global Malmquist indices are executed. For the adjacent Malmquist, the FEAR R-package [WILS13] was used, while an R-script was written for the calculation of the Global Malmquist Indices (Refer to Appendix D, Section 1).

### 7.3.2 Rolling Window Analysis

To provide additional trend analysis and to verify the results of the Malmquist analysis, rolling window analysis, another trend analysis technique commonly used with DEA, has also been applied to the growth model. Since both trend analysis techniques should indicate similar directions of growth, their comparison can provide insight into their applicability to the growth model. Before the rolling window analysis can be carried out, the ideal number of time periods included in each window must be calculated. This can be done using the generally accepted
formula: \( p = k + \frac{1}{2} \); where \( p \) is the number of time periods in each window and \( k \) is to total number of periods in the data set. Given there are 6 time periods that can be used in 5 distinct growth models, the ideal number of periods was calculated to be 3. Once this was determined the provided data set was reshaped and the window analysis was carried out. The application and results of this methodology is discussed in detail below. The R-scripts written for this section are available in Appendix D, Section 1.

7.4 Detailed Application

In order to evaluate the proposed methodology, the developed formulation was applied to an extensive panel data set obtained from a large Canadian Banking Institution. Subsequent to being cleaned, the data set consisted of over 1000 DMUs over 6 years. This section provides a step by step account of the analysis that was carried out including sensitivity analysis, variable selection, validation for the DEA model chosen and the complete application to the data set and the trend analysis techniques. The results of all model executions are evaluated and thoroughly discussed to effectively validate the model.

7.4.1 Variable Selection and Sensitivity Analysis

Variable selection for this application relied heavily on the availability of variables over an extended period of time. Although extensive, the Canadian Banking data set only included customer count, product count and total funds held over several time periods. Consequently, these three variables, along with employment data obtained from other sources, were used to build a growth model focused on production efficiency. Both correlation analysis and the efficiency contribution method were applied to these variables which suggested that although the variables are related they are all integral to the results of the analysis. Moreover, from a managerial standpoint keeping all the variables in the model provides increased credibility and more comprehensive results and target objectives without having a large effect on the discriminatory power of the model. It should be noted that FTEs (t) originally included as an output was removed on the basis of the PCA results as it had no impact on the efficiency distribution of the DEA results. This may not be the case for all growth analyses; variable selection must be performed on a case by case basis.

In addition to the base variables, environmental variables were included in the model to account for local demographic characteristics and local competition. In order to determine which variable combination should be used in the final analysis as well as determine the robustness of
the proposed model, a thorough principal component analysis (PCA) was performed. The complete list of variables that were considered is listed below:

- Competitive Indices: Number of Competitive branches in a 25KM, 10KM, 5KM and 1KM radius
- Population Indices: Local Population (t+1), Local Population (t), Difference in Local Population from (t to t+1)
- Income Indices: Local Average Household Income(t+1), Local Average Household Income (t), Difference in Local Average Household Income from (t to t+1)

To begin the analysis, each competitive index was added to the base model one at a time and the PCA analysis was carried out. If the efficiency distribution changed, that variable was considered for the final model. Of the four competitive indices tested, the 10KM and 5KM index were found to be significant. The efficiency distributions obtained through the inclusion of each of these indices were not found to be significantly different from one another, thus the 5KM index was arbitrarily chosen. It should be noted that the use of clustered data may warrant the inclusion of different competitive indices. For example, it may be more pertinent to use the 25KM radius for rural branches while the 1KM radius may suffice for branches located in large metropolises.

The income and population indices were then evaluated. Originally it was thought that these indices should be included for both periods ‘t’ and ‘t+1’, however the results of PCA determined that the efficiency distribution of the growth model including period ‘t+1’ alone was not significantly different from that of the model containing both periods. In follows that the environmental indices were only included for period ‘t+1’. Subsequently, different combinations of the population and income indices and differences were used in the model and their efficiency distributions compared. It was found that all fourth models (1.Population and Income, 2.ΔPopulation and ΔIncome, 3.ΔPopulation and Income, 4.Population and ΔIncome) produced efficiency distributions that were not significantly different from one another. Each had a range of efficiencies between approximately 0.80-1 and identified between 95 and 123 efficient units, indicating that their discriminatory power was not notably different. In order to maintain simplicity, the straight income and population indices were chosen. The final model used for this application is provided in Table 7.2.
Aside from aiding in the variable selection process, performing PCA also provided insight into the robustness of the model. Given the similarity of the results obtained when evaluating the population and income indices, it is evident that the model is able to consistently measure the growth efficiency of DMUs even when variables are replaced with proxies.

It should also be noted that the use of longer lags (i.e. using time ‘t’ and ‘t+2’ etc.) was contemplated due to the possible existence of a time lag between the initial acquisition cost of a client and the time required to recoup these costs. However, it was determined that lags were not appropriate for this study for several reasons. First and foremost, the study was limited by data availability. Since the data set only contained 5 years of data, extending the lag period would decrease the number of feasible runs of the model. Moreover, it was found that the average acquisition cost for a new retail banking customer is approximately $200-300 while the average customer contribution is approximately $300 per year and the break-even point occurs when the client holds approximately $9000 in their chequing account. [INSI14], [OPTI14] Furthermore, it has been found that once a customer is acquired, the probability of cross selling increases significantly thus increasing potential returns. [BAIN13] Consequently, determining the time required to recoup sunk customer costs requires an extensive amount of customer information to determine the appropriate break-even point for each branch. Lastly, the ultimate goal of the model was to measure a branch’s ability to grow its customer base, funds and clients from one year to the next in order to provide a growth performance efficiency metric. For this application, the use of two consecutive years of data is most appropriate. However, the methodology of incorporating non-contemporaneous data can be modified and executed for applications that require extended lag periods.
7.4.2 **DEA Model Selection**

To ensure the integrity and validity of this analysis, careful consideration must be given to the selection of the DEA model. Given the nature of our inputs and outputs along with the goal of the model to maximize products in year ‘t+1’ an output-oriented model was chosen. Next, several model types were deliberated to determine which could best deal with the scaling issues associated with modeling growth. A summary of this deliberation is provided in Table 7.3.

**Table 7.3: Model Comparison**

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Attributes/Disadvantages</th>
<th>Chosen?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRS</td>
<td>- Simple to execute and easy interpretation of results</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>- One of the most widely used DEA models; very commonly applied to bank branch data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Makes strict assumptions about the RTS and may be too rigid when dealing with growth</td>
<td></td>
</tr>
<tr>
<td>VRS</td>
<td>- Simple to execute and easy interpretation of results</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>- One of the most widely used DEA models; very commonly applied to bank branch data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- May provide results that are too conservative for considering Growth, Includes IRS and DRS units which may not be in line with Bank’s objective and may not be realistic in terms of target objectives</td>
<td></td>
</tr>
<tr>
<td>SBM</td>
<td>- Does not require specification of orientation</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>- Is a Non-oriented model; output-oriented model is desirable for objective of this study</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Added complexity in interpretation of results</td>
<td></td>
</tr>
<tr>
<td>Selective Proportionality</td>
<td>- Allows for the combination of CRS and VRS relationships between variables</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>- Requires that variables be separated into a CRS group and a VRS group, where no overlap can exist. Does not allow variables to have CRS relationships with some variables and VRS with others</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- A lot of added complexity without the RTS flexibility necessary to truly define the relationships within the growth model</td>
<td></td>
</tr>
</tbody>
</table>

Ultimately, for the purpose of this study, it was decided that CRS and VRS models would be employed as they are simple to execute and provide straightforward results. Moreover, they are widely accepted models for use in bank branch studies. Ideally, a selective proportionality model would be used for this analysis; however, the technology necessary to accurately define the variable relationships that exist within the growth model does not exist at this time. Current selective proportionality models would add a large amount of complexity without a proportional payoff. (It should be noted that the developed methodology can be used in conjunction with any DEA model. It follows that one should carefully consider their intended application and their desired result before choosing a model.)
7.4.3 **APPLICATION TO FULL DATA SET**

Once the formulation was finalized, it was applied to the complete Canadian data set for both CRS and VRS models. A summary of the results for each model and year are provided in Table 7.4. The efficiency distributions for each year are provided in Appendix D, Section 2. (Please note that each distribution has a unique scale.)

<table>
<thead>
<tr>
<th>Table 7.4: Complete Data Set- Results Summary</th>
<th>2008-2009</th>
<th>2009-2010</th>
<th>2010-2011</th>
<th>2011-2012</th>
<th>2012-2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRS Mean Efficiency</td>
<td>0.8857</td>
<td>0.8979</td>
<td>0.9470</td>
<td>0.9275</td>
<td>0.9211</td>
</tr>
<tr>
<td>CRS Min. Efficiency</td>
<td>0.5178</td>
<td>0.7083</td>
<td>0.7629</td>
<td>0.7631</td>
<td>0.7677</td>
</tr>
<tr>
<td>CRS # Efficient Units</td>
<td>99</td>
<td>85</td>
<td>104</td>
<td>103</td>
<td>104</td>
</tr>
<tr>
<td>CRS % Efficient Units</td>
<td>9.56</td>
<td>8.20</td>
<td>10.0</td>
<td>9.94</td>
<td>10.0</td>
</tr>
<tr>
<td>VRS Mean Efficiency</td>
<td>0.9145</td>
<td>0.9134</td>
<td>0.9571</td>
<td>0.9425</td>
<td>0.9381</td>
</tr>
<tr>
<td>VRS Min. Efficiency</td>
<td>0.6191</td>
<td>0.7293</td>
<td>0.8294</td>
<td>0.7940</td>
<td>0.8067</td>
</tr>
<tr>
<td>VRS # Efficient Units</td>
<td>180</td>
<td>144</td>
<td>172</td>
<td>182</td>
<td>176</td>
</tr>
<tr>
<td>VRS % Efficient Units</td>
<td>17.37</td>
<td>13.90</td>
<td>16.60</td>
<td>17.57</td>
<td>16.99</td>
</tr>
<tr>
<td>VRS Scale Efficiency</td>
<td>0.9685</td>
<td>0.9830</td>
<td>0.9895</td>
<td>0.9840</td>
<td>0.9818</td>
</tr>
</tbody>
</table>

The efficiency distributions achieved with this model were similar to those commonly reported in the literature, having a near normal distribution. [GOLA99] Moreover, the efficiency ranges were consistent with those obtained through traditional bank branch analysis models. The number of efficient units identified was consistently lower than the 20-30% generally reported from production models. This suggests that this model has a significantly better discriminatory power and is able to identify units that are growing most efficiently given their local environment. The results of the average scale efficiency calculation provide additional insights into banking operations suggesting that the bank operates under approximately 98% scale efficiency. This supports the generally accepted notion that Canadian Bank branches operate under predominantly CRS. ([SCHA97], [WU06a])

Investigation into individual firms further supports the results of the growth model. Following the definition introduced in the previous chapter, a small number of ENG units were found on the efficient frontier. These DMUs were found to largely consist of units that experienced losses in either their funds or customers from one year to the next but not both. This occurs due to the un-restricted DEA model assigning zero or very small multipliers to variables that experience loss and larger multipliers to the variables which experience growth. The few DMUs which experience loss in both funds and customers were found to be extreme outliers.
operating in downtown cores or extreme demographic regions. These misspecifications are unique cases which did not deter from the overall performance of the growth model in properly identifying growth efficiency. However, to deal with these ENG units, a combined methodology is introduced in Section 7.5.

Through the application of the growth model to the entire Canadian Banking data set, it can be seen that not only does this model effectively provide discriminating growth efficiency scores, peer groups and target objectives for each bank branch, it also provides better discrimination than traditional DEA models. Moreover, the model was found to better appreciate the environmental factors related to branch growth. In summation, this model was found to be highly effective at evaluating growth efficiency and was able to provide a comprehensive growth analysis of the branch network.

7.4.4 DMU Categorization

In order to provide a more comprehensive evaluation of a DMU’s performance, the results of the growth model were compared to the results of a traditional production model whose inputs and outputs are provided in Table 7.5. This offers a more detailed characterization of the DMU that can be used to identify whether the units should put more focus on improving its growth efficiency, production efficiency or both.

**Table 7.5: Traditional Production Model**

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>Products</td>
</tr>
<tr>
<td>- FTEs</td>
<td>- Amount of Funds</td>
</tr>
<tr>
<td>Indices</td>
<td>- Number of Products</td>
</tr>
<tr>
<td>- Population Index</td>
<td>- Number of Customers</td>
</tr>
<tr>
<td>- Income Index</td>
<td>Competitive Index (5KM)</td>
</tr>
</tbody>
</table>

This methodology was executed for 2011 and compared to the 2010-2011 growth model. The results can be depicted in several ways depending on the manager’s interests and objectives. For example Figure 7.2 illustrates central cross hairs while Figure 7.3 places 25% of the units in the top right “best performers” quadrant. Similarly, one can identify a certain percentage of worst performers that need immediate attention by decreasing the size of the bottom left quadrant. Both Figure 7.2 and 7.3 depict CCR results.
Plotting growth efficiency against production efficiency provides a heuristic means of characterizing each DMU. DMUs in the top right quadrant can be considered high achievers in
both growth and production. DMUs in the top left quadrant can be considered ‘sleepers’; they are performing relatively well in terms of growing their assets, but are not working to their production capacity. DMUs located in the bottom right quadrant are producing efficiently but have untapped market potential that could lead to greater growth. Finally, DMUs in the bottom left quadrant are troubled units. The bank should focus its attention on these units as they are operating inefficiently in all senses.

It should be noted that in all cases when a DMU was found to be efficient in the production model, it was also efficient in the growth model. The points highlighted in red on both graphs are, in reality, many points sitting on top of one another. These results are not at all surprising as being the most efficient producer in a set of DMUs means these units have the greatest growth potential.

### 7.4.5 Clustered Data

To provide additional insight into the growth efficiency of each branch, a local analysis was also performed. This can provide more focused target objectives and peer groups for inefficient DMUs. Moreover, the impact of economic trends and local characteristics can be examined further through this analysis.

To begin, the within group sum of squares graph was inspected to determine the appropriate number of clusters to be between 7 and 8. (Refer to Section 5.3.3, Figure 5.1) To ensure that each cluster had a sufficient number of DMUs, 7 clusters were used. K-means was then used to determine the DMU groupings for each of the 7 clusters based on the population, income and competitive indices. The number of DMUs in each cluster is outlined in Table 7.6.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Type</td>
<td>High Income Metropolitan</td>
<td>70% Urban, 30% Rural</td>
<td>High Income Suburban</td>
<td>Mid Income Metropolitan</td>
<td>Low Income Rural</td>
<td>Mid Income Suburban</td>
<td>Mid Income Rural</td>
</tr>
<tr>
<td>Population</td>
<td>High</td>
<td>Mid</td>
<td>Mid/High</td>
<td>High</td>
<td>Low</td>
<td>Mid</td>
<td>Low</td>
</tr>
<tr>
<td>Income</td>
<td>High</td>
<td>High</td>
<td>Very High</td>
<td>Mid/High</td>
<td>Low</td>
<td>Mid</td>
<td>Mid</td>
</tr>
<tr>
<td># DMUs</td>
<td>38</td>
<td>113</td>
<td>51</td>
<td>36</td>
<td>280</td>
<td>238</td>
<td>280</td>
</tr>
</tbody>
</table>

These clusters were then used to determine whether any additional trends could be gleaned from the global analysis. The summary of the global results by cluster are provided in Table 7.7. It can be seen that the high income metropolitan area has the highest growth efficiency, while the mid income suburban area has the lowest. It is likely that the growth
experienced by the metropolitan areas can be attributed to the fast paced, dynamic nature of
downtown cores where there is a steady influx of new potential clients. The suburban area’s low
growth efficiencies are most likely caused by the more stagnant suburban environment which
does not experience significant changes in population from one year to the next. Consequently,
there are fewer new potential clients for the branch to attract.

Table 7.7: Global Analysis Cluster Results

<table>
<thead>
<tr>
<th>Cluster Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>0.995</td>
<td>0.960</td>
<td>0.943</td>
<td>0.952</td>
<td>0.949</td>
<td>0.939</td>
<td>0.940</td>
</tr>
<tr>
<td>Min</td>
<td>0.97</td>
<td>0.872</td>
<td>0.886</td>
<td>0.856</td>
<td>0.763</td>
<td>0.817</td>
<td>0.821</td>
</tr>
<tr>
<td>%Efficient</td>
<td>0.74</td>
<td>0.124</td>
<td>0.118</td>
<td>0.194</td>
<td>0.064</td>
<td>0.064</td>
<td>0.064</td>
</tr>
</tbody>
</table>

A local analysis was also performed for each cluster. The results of these analyses are
summarized in Table 7.8. It should be noted that while the global analysis results include the use
of indices, the local analyses do not. Since the environmental variables were already used to
determine the DMU clusters, they were not included in the inputs and outputs of the local
analyses. (i.e. The growth model was run without indices for each local analysis.) This results in
improved discrimination and lower average efficiency scores.

Table 7.8: Local Analysis Results-CRS

<table>
<thead>
<tr>
<th>Cluster Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>0.980</td>
<td>0.969</td>
<td>0.931</td>
<td>0.925</td>
<td>0.953</td>
<td>0.914</td>
<td>0.967</td>
</tr>
<tr>
<td>Min</td>
<td>0.943</td>
<td>0.886</td>
<td>0.830</td>
<td>0.795</td>
<td>0.884</td>
<td>0.779</td>
<td>0.926</td>
</tr>
<tr>
<td>%Efficient</td>
<td>0.342</td>
<td>0.221</td>
<td>0.216</td>
<td>0.306</td>
<td>0.121</td>
<td>0.113</td>
<td>0.104</td>
</tr>
</tbody>
</table>

Although DEA runs are not directly comparable, especially when the models’ inputs and
outputs differ, some interesting information can be obtained through the comparison of the local
and global clustered results. For certain clusters, the local analysis’ average efficiency score is
lower than that of the global analysis. Upon further investigation, it is found that certain DMUs
that were deemed to be efficient in the global analysis were no longer efficient in the without
indices local analysis. With the indices removed, the model no longer captures the subtle
differences in income, population and external competition between branch environments. Since
each cluster contains units that are functioning in very similar environments, the differences
caused by the removal of the indices is not a concern. On the contrary, the local analysis provides a means of accounting for environment, while improving discrimination.

When comparing each cluster, it can be seen that certain clusters have lower average efficiency scores and certain clusters have larger differences in efficiencies from year to year. This provides additional information as to which local demographic ranges result in branches that are more susceptible to operational inefficiencies as well as which demographic ranges were most affected during the recession. Much like the results of the global model, the Metropolitan area had the highest growth efficiency while the mid-income suburban area had the lowest. This information is invaluable in developing solutions to increase growth in these underperforming regions and in understanding how the role that demographics play in achieving growth within the branches.

In conclusion, the cluster analysis provided many important insights into the growth efficiency of branches located in different demographic regions. The information obtained from this analysis is crucial in obtaining a better understanding how demographic characteristics effect branch grow and is invaluable for creating tailored improvement plans. Moreover, this application further validated the usefulness and applicability of the proposed growth model formulation.

7.4.6 Applying Malmquist Indices

To provide additional information pertaining to growth trends as well as a detailed look into frontier shifts, both Adjacent and Global Malmquist indices were calculated.

Adjacent Malmquist Model

Due to the dynamic nature of the growth model, the interpretation of the derived Malmquist indices must be carefully considered. The Malmquist index itself must be interpreted as an indicator of the change in growth efficiency of the branches from one year to the next. It follows that the efficiency catch up component is the DMU’s improvement in growth while the frontier shift denotes the change in the branch network’s growth potential from one year to the next.

The results of the Adjacent CCR Malmquist model (Table 7.9) provide additional insights into the changes in growth from one fiscal period to the next. The overall Malmquist indices tend to suggest that the Bank maintains a fairly constant growth rate from one year to the next with little fluctuation. This supports the Bank’s claim that their customer churn is netting zero from
one year to year. Additional information about how the branches are maintaining this growth is visible through the decomposition of the index. The frontier shift component (i.e. technology change) is moving forward for the 2008-2009 to 2009-2010 comparison as well as for the 2009-2010 to 2010-2011. The growth begins to slow in the final columns with Frontier Shift components of approximately 1, meaning no frontier advancement from year to year. When looking at the efficiency change components in columns one and two, the average DMU is lagging behind the frontier slightly, while the catch up for columns three and four show on average the efficiency range of the DMUs is becoming smaller as they all approach the frontier.

Table 7.9: Adjacent Malmquist Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Malmquist Index</td>
<td>1.000</td>
<td>0.991</td>
<td>1.00238</td>
<td>1.005</td>
</tr>
<tr>
<td>Efficiency Catch Up</td>
<td>0.987</td>
<td>0.954</td>
<td>1.019301</td>
<td>1.011</td>
</tr>
<tr>
<td>Frontier Shift</td>
<td>1.014</td>
<td>1.028</td>
<td>0.983514</td>
<td>0.994</td>
</tr>
</tbody>
</table>

The results of this application support those obtained from the efficiency analysis and demonstrate that the proposed growth model is also applicable to growth trend analysis. The Malmquist indices are able to clearly present the branch network’s improvements in growth efficiency from one year to the next, which is highly desirable from a managerial stand point.

Global Malmquist Model

To obtain a sense of the overall growth efficiency changes over time the Global Malmquist indices were calculated for each branch. A summary of the CCR results is provided in Table 7.10.

Table 7.10: Global Malmquist Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Malmquist Index</td>
<td>1.005</td>
<td>1.000</td>
<td>0.997</td>
<td>0.996</td>
</tr>
<tr>
<td>Efficiency Catch Up</td>
<td>0.987</td>
<td>0.947</td>
<td>1.0218</td>
<td>1.008</td>
</tr>
<tr>
<td>Frontier Shift</td>
<td>1.020</td>
<td>1.0577</td>
<td>0.976</td>
<td>0.989</td>
</tr>
</tbody>
</table>

In terms of its decomposed components, the global Malmquist index produced very similar results to those obtained with the adjacent indices. The differences between the efficiency catch up and frontier shift components of the first two columns compared to the third and fourth column are slightly more pronounced but follow the same pattern. These larger component
differences result in some interesting changes in the calculated Malmquist indices. Contrary to what was expected, the global Malmquist indices indicate that branch growth is steadily decreasing from year to year. From here the growth steadily reaches a plateau producing Malmquist indices of approximately 1. A reasonable explanation for this is that, as indicated by bank management, the overall growth of clients from one year to the next is net zero resulting in no change in the growth over that time.

It should be noted that given the unique model structure of the growth efficiency model, the calculation of the global Malmquist index will have certain outputs that are related to performance in time ‘t’ that will also enter into the model as inputs related to performance in time ‘t+1’. This unique structure may make the index more sensitive to random events. For example, if a given branch experiences an unexpected decrease in funds just before the end of period ‘t’, and shortly after, in the beginning of period ‘t+1’, realizes an unexpected increase the branch will be said to have poor growth performance in time ‘t’ and good performance in time ‘t+1’. The high growth rate in time ‘t+1’ may dominate other more stable branches thus skewing the global frontier. It follows that the global frontier used to calculate the global Malmquist index is sensitive to inflated maximal growth rates resulting from random effects occurring close to the time when data is collected. Although this sensitivity exists, it does not appear to have affects the results of the models run herein.

Understanding how branch characteristics affect the frontier shift and the DMU’s ability to improve its growth standing is vital for creating realistic and achievable improvement plans. Consequently, the relationship between the growth trends and branch characteristics was investigated through clustering the Malmquist results. Using the same clusters as defined in Table 7.6, the results from the Adjacent Malmquist analysis are summarized by cluster in Table 7.11. (Note this table summarizes the results comparing the 2010-2011 frontier to the 2011-2012 frontier.)
Table 7.11: Malmquist Cluster Comparison: 2010-2011 Frontier to 2011-2012 Frontier

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Malmquist Index</th>
<th>Efficiency Catch Up</th>
<th>Frontier Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : High Income Metropolitan</td>
<td>1.004</td>
<td>1.045</td>
<td>0.940</td>
</tr>
<tr>
<td>2: 70% Urban, 30% Rural</td>
<td>0.992</td>
<td>1.065</td>
<td>0.939</td>
</tr>
<tr>
<td>3: High Income Suburban</td>
<td>1.014</td>
<td>1.089</td>
<td>0.929</td>
</tr>
<tr>
<td>4: Mid Income Metropolitan</td>
<td>0.932</td>
<td>1.040</td>
<td>0.898</td>
</tr>
<tr>
<td>5: Low Income Rural</td>
<td>0.996</td>
<td>1.024</td>
<td>0.971</td>
</tr>
<tr>
<td>6: Mid Income Suburban</td>
<td>1.000</td>
<td>1.023</td>
<td>0.958</td>
</tr>
<tr>
<td>7: Mid Income Rural</td>
<td>0.995</td>
<td>1.014</td>
<td>0.981</td>
</tr>
</tbody>
</table>

The results show that all demographic regions have an efficiency catch up of greater than one and a frontier shift of less than one. This means that the efficiency distribution is becoming less broad and inefficient units are performing more similarly to the efficient units. However, the frontier is regressing from one year to the next. This suggests that although inefficient branches are performing better from one year to the next relative to their peers, the branches, overall, are performing worse in terms of branch growth from year to year. Ultimately, this has resulted in the majority of the demographic regions having average Malmquist scores of approximately one.

7.4.7 Rolling Window Analysis

To conclude the growth trend portion of this application, a rolling window analysis was performed on the data set. The results of this analysis followed the trends previously observed in from the global and adjacent Malmquist analyses. The average efficiency scores obtained for window and year are provided in Table 7.12.

Table 7.12: Rolling Window Analysis- Average Efficiency

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Window 1</td>
<td>0.865</td>
<td>0.862</td>
<td>0.863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Window 2</td>
<td></td>
<td>0.872</td>
<td>0.871</td>
<td>0.873</td>
<td></td>
</tr>
<tr>
<td>Window 3</td>
<td></td>
<td>0.971</td>
<td>0.968</td>
<td>0.973</td>
<td></td>
</tr>
</tbody>
</table>

The results show that there are insignificant differences between the growth efficiencies from one year to the next within each window. Looking at Window 1, for example, we see that the average efficiency score is essentially the same for all periods. This suggests that there is little to no change in the growth efficiencies over time. This pattern continues in Windows 2 and 3. This not surprising as banks tend to highly centralise and control their branch operations so changes are very slow.
7.5 **COMBINED METHODOLOGY**

To complete this investigation of growth efficiency in bank branches, both methodologies were combined to provide a restricted growth efficiency model. To execute this model, the mathematical formulation of the restricted model was applied to the growth model methodology developed above. Table 7.13 provides a comparison of the results obtained from the growth model and the combined model when applied to the Canadian dataset.

<table>
<thead>
<tr>
<th></th>
<th>CRS Growth Model</th>
<th>CRS Combined Model</th>
<th>VRS Growth Model</th>
<th>VRS Combined Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
<td>0.947</td>
<td>0.954</td>
<td>0.957</td>
<td>0.963</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>0.763</td>
<td>0.7585</td>
<td>0.829</td>
<td>0.866</td>
</tr>
<tr>
<td># of Efficient Units</td>
<td>104</td>
<td>140</td>
<td>172</td>
<td>202</td>
</tr>
<tr>
<td># units &gt;1 Efficiency</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>18</td>
</tr>
<tr>
<td>#ENG units</td>
<td>15</td>
<td>-</td>
<td>23</td>
<td>-</td>
</tr>
<tr>
<td># Infeasible Units</td>
<td>-</td>
<td>8</td>
<td>-</td>
<td>15</td>
</tr>
</tbody>
</table>

Due to the strict definition of growth that was used to determine whether each DMU had negative or positive annual growth, there were a small number of ENG DMUs identified when the Growth model was employed. Upon further investigation, it was noted that the majority of these units had negative growth in either customers or funds but not both. The remaining units were all found to be anomalous outlier units or units that functions in extreme demographics (e.g., very low income and very low population, etc.) For example, in the CRS case 13 of the 15 identified ENG units experienced negative growth in either the number of customers or the amount of funds. The remaining two ENG units consisted of the very large central branch in Calgary’s downtown core and a branch that operated in a very low population and low income area. These two branches were identified by the growth model as being efficient since they were the sole units sitting in either region on the frontier. This observation not only validates the growth model’s ability to properly identify growing DMUs but it also highlights the usefulness of combining the models for outlier identification and to impose stricter growth restrictions.

From Table 7.13 it can also be seen that the number of efficient units increases when the restriction is applied. This is due to the fact that once the ENG units are removed, the efficient frontier drops to a larger number of DMUs that all share the frontier and are positive growth best performers. Consequently, the combined methodology is able to provide an improved growth performance measure and peer groups that consist of exclusively positive growth units.
The combined methodology was able to further illustrate the usefulness and validity of the methodologies developed in this work as well as to provide a robust and comprehensive measure of bank branch growth efficiency.

7.6 CONCLUSIONS

This study has successfully demonstrated the usefulness and applicability of the newly developed growth model through the validation and analysis of its results. The model itself was able to properly identify the growth efficiencies of bank branches and only misidentified a small number of DMUs due to the unrestricted nature of the DEA model used. The growth model was also able to provide inferences about the overall growth efficiency of the branch network at a given time. Moreover, the application of cluster analysis was able to provide more tailored growth target objective and peer groups as well as additional insights into which demographic regions are more prone to housing negative growth units. Using slightly altered result interpretations, growth trends were successfully analyzed through the use of Malmquist indices and rolling window analysis. The information that can be obtained from this approach is valuable as they identify troubled branches and demographic regions of concern, understanding growth trends, and developing realistic growth improvement strategies.

The Combined methodology which employed both the growth model and the restricted growth model introduced in Chapter 6 provided a final robust model that was able to effectively measure the growth efficiency of each branch while ensuring the frontier was free of any negative growth units. Moreover, the combined technique was able to further validate both methodologies developed in this work and provide a new, unique and comprehensive solution to measuring the growth of bank branches.

Not only was the growth model highly successful in analyzing all facets of branch growth, this study also provides a substantial contribution to the field of DEA by providing a new method of model formulation that incorporates multiple years of data into one model, consequently redefining the concepts of an input and output.
Chapter 8: Conclusions and Recommendations

This chapter presents a brief synopsis of the research, along with the conclusions drawn from the examination and analysis of the results. Directions for future work follows.

8.1 Restricted Model

The restricted growth model developed during this study imposes an additional constraint on the non-negative intensity variable for all DMUs that have experienced negative growth from one year to the next. All DMUs to which this constraint is applied cannot fall within the MPPS and therefore cannot be referenced by inefficient units. This ensures that all peer groups consist of positive growth units and all operational target objectives are associated with producing positive growth.

This approach was tested and validated through its application to three distinct data sets; one simulated, one provided by a Canadian Bank and one provided by a Turkish Bank. Through these applications all infeasibilities were identified and the methodology’s usefulness and validity were verified. It was found that, although there are conditions that result in infeasible solutions, the integrity of the model and the usefulness of the results are not diminished by these infeasibilities. Inefficient units, as well as most ENG units, were provided more realistic peer groups and the efficiency scores of inefficient units were corrected to properly account for growth. Moreover, this methodology was able to identify and removed negative growth outlier units from the efficient frontier as well as identify demographic regions that are susceptible to negative growth units. It should be noted, that the definition of negative growth that was used for the restriction in these applications was fairly harsh and can easily be modified for this or other applications where one wishes to restrict efficiencies based on a specific external variable.

As a last form of validation, the application of other restricted DEA models was investigated. As cone ratios could not be obtained for this data set, the EXFA model was applied; however it was unable to achieve the objectives set for the restricted growth model. It follows that the restricted growth model provides a significant theoretical contribution to the field of DEA as it bridges the gap between traditional DEA models and proper growth accounting.
8.2 Growth Model

The Growth model developed in this work captures the dynamic nature of growth through the inclusion of non-contemporaneous data into a DEA model in order to properly assess the grow efficiency of bank branches. This study has successfully validated the proposed technology as well as demonstrated the usefulness of its application. With the exception of a small number of ENG units, the model was able to effectively measure the growth efficiency of each DMU as well as identify demographic regions which are more susceptible to negative growth branches. Through the application of Malmquist indices, growth trends were successfully identified. Together the results of this study can be used to develop realistic performance objectives and improvement strategies for each branch and/or demographic region.

In addition to the real world application of the results in this study, this model provides several significant contributions to DEA theory. DEA Bank Branch studies are less common than institutional analyses due to the lack of proprietary bank branch data. Through a significant effort in data collection, a comprehensive data set was compiled which included demographic and competitive indices. Furthermore, until this time no DEA literature focused directly on identifying the growth potential within bank branches. This model provides a method to specifically measure the growth efficiency of bank branch networks. Finally, this approach provides a new method of model formulation that employs non-contemporaneous data within one DEA model. This new definition of what comprises an "input" or an "output" results in a ‘dynamic’ DEA model that is able to capture branch growth.

In summation, this work advanced both the theoretical and methodological components of DEA providing a comprehensive and realistic analysis of bank branch growth efficiency.

8.3 Combined Methodologies

By combining the two approaches, a robust growth efficiency model was developed to assess the growth performance of bank branch networks. The model was able to properly assess each branch’s ability to grow customers, products and funds from one year to the next while ensuring that any negative growth units were restricted from being in the MPPS. This approach also provided a means of removing outlier units from the efficient frontier and helped to identify demographic regions where negative growth branches were common.
8.4 Methodological Validation

Throughout this work it was demonstrated that each of the developed methodologies met its objective, providing a robust means of accounting for growth within a modified DEA framework. However, further external forms of validation can be useful in determining whether the methodologies fall in line with the Bank’s in-house performance analysis systems and whether they provide successful growth improvement strategies.

Banks employs several techniques to assess the performance of their branch networks in order to determine who the best and worst performers are. Comparing the bank’s internal results with those obtained through the application of the developed methodologies provides a means of assessing whether the methodologies are in line with the views of Bank management. Similar results would provide further validation of the models, although discrepancies should not be disregarded as they often times highlight unique cases that should be investigated further. Unfortunately, bank performance metrics were not available for either bank examined in this study.

Another form of external validation would be to apply the target objectives obtained through this analysis to each inefficient branch and track the branches’ growth efficiencies over a sufficient amount of time. This would provide invaluable insight into whether the objectives provided by these methodologies are, in fact, useful for improving branch growth efficiency. This process would require a large amount of co-operation from the bank as well as ample time to observe the impact of the changes on the branch’s efficiency.
Chapter 9: Directions for Future Work

Throughout the course of this study, several other potential research avenues related to branch growth and efficiency has come to light. This section provides possible directions for future work in this exciting area of study.

9.1 Growth Efficiency and Age Demographics

In this study, income, population and competitive indices were successfully employed to incorporate important environmental factors into the growth analyses. It may be insightful to extend this work to incorporate the age profiles of each branch’s customer base. These profiles may have a significant impact on branch operations, product demand, and the branch’s overall ability to grow. Additionally, the age distribution may have a large impact on the effectiveness of a Bank’s improvement strategies. (eg. should they be marketing to older or younger clientele?)

In order to perform a thorough analysis of this hypothesis, a large amount of customer demographic data would be required from the Bank. Due to privacy concerns banks are generally reluctant to provide such data and thus obtaining it may be very difficult. Alternately, Statistics Canada data may prove to be a sufficient proxy. Detailed age data would need to be collected for each dissemination area and then matched to each branch, as performed for the income and population indices.

The method of including the age profiles into the DEA analysis would also need to be carefully considered. The use of average age may not accurately depict the true nature of the population while incorporating several indices for age ranges may have a negative impact on the dimensionality of the model.

9.2 Extend Restricted Model to Deal with Infeasibilities

Although the infeasible solutions associated with the restricted model are not detrimental to the model’s validity and applicability, it may be useful to investigate ways of dealing with these infeasibilities in order to provide more informative solutions to the currently infeasible ENG units. This will require modifications to the mathematical formulation of the model. One possible avenue of investigation could be the methodologies developed in [COOK09], [CHEN11], and [LEE11].
9.3 **Extended Facet Analysis (EXFA): Non-negative Growth Facets**

Extended Facet Analysis, introduced in Section 3.8.6 and applied in Section 6.6, allows for the facet structure of the efficient frontier to be restricted. In this study, the EXFA model was applied in order to determine whether restricting the structure of the frontier to fully dimensional facets would have an impact on the presence of ENG units. It would be interesting to extend this analysis by only allowing facets spanned by branches with non-negative growth. This would provide an alternative method to the restricted model introduced herein.

9.4 **Decomposition of Growth Efficiency**

As was mentioned in Chapter 7, the Growth Efficiency Model introduced in this work was based upon, and in fact subsumes a production model due to data availability. It follows that an interesting and possibly useful extension to this model would be to decompose the raw growth efficiency scores into parts that measure the effects of inefficiency stemming from static production and those solely related to the growth of each branch.

9.5 **Sub-Vector DEA**

Scaling of inputs and outputs can pose a large problem when DEA suggests the up scaling or downscaling of a DMU. It is obvious, for example, that if we scale up the unit by a factor of 2, it is most likely not realistically feasible for the growth terms to also double. One way to deal with this issue would be to use selective proportionality. This would allow more precise variable relationships to be incorporated into the model to account for the non-uniform scaling. Although this sounds like a promising approach, the current technology is not able to deal with the complexities of the variable relationships that exist within the period growth model and thus cannot be applied at this time. Another possible solution to this issue is the use of sub-vector DEA which generates technical efficiency measures for a single input (output) or a subset of inputs (outputs) rather than for the entire vector of inputs (outputs). Since the main focus of this study is to identify which branches are growing most efficiently, performing an output oriented sub-vector DEA analysis on the growth related output variables could provide results that fall in line with the objectives of this model while avoiding issues with scaling.

9.6 **Trade Offs**

As previously introduced, there are two main forms of multiplier restrictions: either value based or technology based. Technology based restrictions, also known as trade-offs, ensure that
the radial targets of the inefficient units are technologically realistic thus the efficient measures remain factors of radial improvement. Not only is this technique beneficial in producing meaningful results but it has yet to be applied in bank branch efficiency analyses. Application of this methodology requires that the operations of the bank branch be carefully considered and the input and output trade-offs thoughtfully determined. Inaccurate or unrealistic trade-offs would be detrimental to the success of the models.

9.7 Determining the Feasibility of Multi-Stage Models

Although growth plays a large role in the overall health of a banking institution, it is not the only factor that should be considered when carrying out a performance analysis. In pursuing a more comprehensive look at bank branch performance the feasibility of a multi-stage model which incorporates the Period Growth Model along with other growth models introduced in [LAPL15] should be investigated. The addition of more traditional production, profitability and intermediation models may also provide favourable results.

9.8 Meta Frontiers

Meta frontiers provide the opportunity to compare frontiers from individual subgroups to the overall grouped frontier. Much like the global frontier used in the calculation of the global Malmquist index, a Meta frontier consisting of the DMUs from all time periods can be created. Each year can be treated as a subgroup with its own unique frontier that can then be compared to the Meta frontier. In order to do this successfully, the values of all of the variables measured in currency will be brought to a common base period value of money and the effects of inflation will be accounted for. Once this is completed and the models are run, each DMU can be compared to their respective time period frontier as well as the Meta frontier. This will provide insights into how well a unit is growing with respect to all the units over all time periods. Additionally, it will be interesting to see which units are on the Meta frontier, as they would be the overall best performers.

9.9 Product Mix Efficiency

Another topic of interest is which is loosely related to branch growth is investigating product mix efficiency. In my previous work, [LAPL12] extended in [LAPL15], a Market model was successfully developed to evaluate the market efficiency of the Canadian Bank’s branch network. Given the new availability of location and demographic data, this model can now be
used to determine the most efficient product mixes for different demographic regions. Preliminary work has been completed investigating the relative effect of different products on the efficiency distributions obtained from the model. The results showed that RRIF and Investments balance had little to no effect on the efficiency distributions while deposits and lending were very influential. This methodology could now be applied to different demographic segmentations to maximize efficiency and determine the most desirable product mix for each branch type. This information would be invaluable for bank management and would allow for tailored branch specific marketing and could be used to create more successful incentive schemes.
### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABM</td>
<td>Automated Banking Machine; Self-service machines that allow customers of a financial institution to perform many everyday “banking” tasks, often at locations other than the institution.</td>
</tr>
<tr>
<td>Additive Model</td>
<td>A DEA model which measures efficiency using the slacks only. Involves simultaneous reduction in inputs and increase in outputs.</td>
</tr>
<tr>
<td>Allocative Efficiency</td>
<td>A measure of the ability to combine inputs and outputs in optimal proportions in the presence of market prices.</td>
</tr>
<tr>
<td>Bank Draft</td>
<td>A cheque drawn by a bank against funds deposited in another bank.</td>
</tr>
<tr>
<td>BCC or VRS Model</td>
<td>A DEA model which assumes a variable return to scale relationship between inputs and outputs.</td>
</tr>
<tr>
<td>Categorical Variable</td>
<td>A variable that assigns a DMU to a specific class with predefined discrete values.</td>
</tr>
<tr>
<td>CCR or CRS Model</td>
<td>A DEA model which assumes a constant return to scale relationship between inputs and outputs.</td>
</tr>
<tr>
<td>Combined Model</td>
<td>A DEA model employing the results of three distinct DEA analyses as outputs and a fixed input of 1.</td>
</tr>
<tr>
<td>Correlation</td>
<td>A measure of the strength of the relationship between two coefficient variables. The value lies between +1 and -1, with a score of 0 implying no relationship.</td>
</tr>
<tr>
<td>CRS</td>
<td>Constant Returns to Scale; A measure where a proportionate increase in inputs results in an identical proportionate increase in outputs.</td>
</tr>
<tr>
<td>DEA</td>
<td>Data Envelopment Analysis; A non-parametric, linear programming technique used to measure the relative efficiency of homogenous decision making units.</td>
</tr>
<tr>
<td>DFA</td>
<td>Distribution Free Approach; A parametric frontier approach to performance analysis.</td>
</tr>
<tr>
<td>DMU</td>
<td>Decision Making Unit; A term used to describe a unit being analyzed by DEA, in this case bank branches.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>DRS</td>
<td>Decreasing Returns to Scale; A measure where a proportionate increase in inputs results in a less than proportionate increase in outputs.</td>
</tr>
<tr>
<td>Economies of Scale/Returns to Scale</td>
<td>The changes in output resulting from proportional changes in input.</td>
</tr>
<tr>
<td>EFA</td>
<td>Econometric Frontier Approach; Also referred to as Stochastic Frontier Approach (SFA).</td>
</tr>
<tr>
<td>Efficiency</td>
<td>A measure of the ability to produce outputs using a minimum level of inputs.</td>
</tr>
<tr>
<td>Efficient Frontier</td>
<td>Empirical frontier that represents the “best performance” and consists of the decision making units in the data set that are most efficient at transforming inputs into outputs.</td>
</tr>
<tr>
<td>ENG</td>
<td>Efficient Negative-Growth unit</td>
</tr>
<tr>
<td>FDH</td>
<td>Free Disposal Hull; A non-parametric linear programming technique that constructs a stepwise efficient frontier.</td>
</tr>
<tr>
<td>FTE</td>
<td>Full-Time Equivalent; Total number of full time employees required to complete the task.</td>
</tr>
<tr>
<td>GRS</td>
<td>Generalized Returns-to-Scale denotes a BCC model whose convexity constraint is given a lower and upper bound.</td>
</tr>
<tr>
<td>Index Number</td>
<td>A statistic that measures the change in weighted variables with respect to a standard or base.</td>
</tr>
<tr>
<td>Input-Oriented Model</td>
<td>A DEA model whose objective is to minimize inputs while keeping outputs constant.</td>
</tr>
<tr>
<td>IRS</td>
<td>Increasing Returns to Scale; A measure where a proportionate increase in inputs results in a more than proportionate increase in outputs.</td>
</tr>
<tr>
<td>Market Efficiency</td>
<td>The extent to which individual branches, given their capacity and available resources, utilize their market potential by maximizing sales.</td>
</tr>
<tr>
<td>Market Model</td>
<td>A DEA model which captures a bank branch’s ability to utilize their market potential to maximize sales using staffing, and market and branch characteristics as inputs and sales as outputs.</td>
</tr>
<tr>
<td>Mix Inefficiency</td>
<td>The non-radial inefficiencies, calculated by dividing the SBM efficiency score by the CCR efficiency score.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>MPSS</td>
<td>Most Productive Scale Size; The point on the efficiency frontier where the maximum average productivity is achieved for a given input/output mix.</td>
</tr>
<tr>
<td>NDRS</td>
<td>Non-Decreasing Returns-to-Scale denotes a BCC model whose convexity constraint is $1 \leq \sum \lambda_j$ resulting in only constant or increasing returns-to-scale.</td>
</tr>
<tr>
<td>NIRS</td>
<td>Non-Increasing Returns-to-Scale denotes a BCC model whose convexity constraint is $1 \geq \sum \lambda_j$ resulting in only constant or decreasing returns-to-scale.</td>
</tr>
<tr>
<td>Non-Discretionary Variable</td>
<td>A variable that is pertinent to the analysis but whose level of use or production is uncontrollable by management (i.e. weather).</td>
</tr>
<tr>
<td>Output-Oriented Model</td>
<td>A DEA model whose objective is to maximize outputs while keeping inputs constant.</td>
</tr>
<tr>
<td>Overall Efficiency</td>
<td>The efficiency measured as a product of technical and allocative efficiency.</td>
</tr>
<tr>
<td>Peer Group</td>
<td>A set of efficient units to which the inefficient has been most directly compared within a DEA analysis.</td>
</tr>
<tr>
<td>Production Function</td>
<td>The function in which outputs are defined as functions of inputs.</td>
</tr>
<tr>
<td>Production Model</td>
<td>A DEA model that captures the business operations of a bank branch using staffing as inputs and transactions as outputs.</td>
</tr>
<tr>
<td>Production Possibility Set</td>
<td>Given the observed data, the set of all possible input/output combinations that could exist.</td>
</tr>
<tr>
<td>Productive Efficiency</td>
<td>The extent to which individual units, given their available resources, are able to produce outputs, in this case transactions.</td>
</tr>
<tr>
<td>Profitability Efficiency</td>
<td>The extent to which individual units are able to maximize revenues while minimizing expenses.</td>
</tr>
<tr>
<td>Profitability Model</td>
<td>A DEA model that captures the business operations of a bank branch using revenues as outputs and branch expenses as inputs.</td>
</tr>
<tr>
<td>PUC</td>
<td>Product Use Count is a term used by the “Bank” to represent the number of products held by a client. The PUC of a branch is the sum of the branch’s client PUCs.</td>
</tr>
<tr>
<td>Relative Efficiency</td>
<td>The measure of actual performance of a unit relative to the best observed performance of other units in the set.</td>
</tr>
<tr>
<td><strong>SBM</strong></td>
<td>Slacks Based Model.</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Scale Efficiency</strong></td>
<td>Efficiency that indicates whether a unit is operating at its optimal size. This measure is calculated as a ratio of CCR to BCC efficiency.</td>
</tr>
<tr>
<td><strong>SFA</strong></td>
<td>Stochastic Frontier Approach; A parametric frontier approach to performance analysis.</td>
</tr>
<tr>
<td><strong>Share of Wallet (SOW)</strong></td>
<td>The percentage (share) of a customer’s financial products (wallet) that are with the firm in question.</td>
</tr>
<tr>
<td><strong>Slack Variable</strong></td>
<td>Slacks represent the over-utilization of inputs or the under-production of outputs in a DEA evaluation.</td>
</tr>
<tr>
<td><strong>Target</strong></td>
<td>The value of inputs and outputs that would result in an inefficient unit becoming efficient.</td>
</tr>
<tr>
<td><strong>Technical Efficiency</strong></td>
<td>The measure of the ability of a unit to produce the maximum output for a given set of inputs.</td>
</tr>
<tr>
<td><strong>TFA</strong></td>
<td>Thick Frontier Approach; A parametric frontier approach to performance analysis.</td>
</tr>
<tr>
<td><strong>Theoretical Frontier</strong></td>
<td>The frontier of best possible production that is not necessarily based on the performance of the observed units.</td>
</tr>
<tr>
<td><strong>VRS</strong></td>
<td>Variable Returns to Scale; A measure where a proportionate increase in inputs could result in a proportionate increase or decrease in outputs.</td>
</tr>
<tr>
<td><strong>Weight/Multipliers</strong></td>
<td>The coefficients applied to the input and output variables.</td>
</tr>
</tbody>
</table>
References


[BAOT05] The Banks Association of Turkey, “Financial Sector and Banking System in Turkey”, Report Issued by The Banks Association of Turkey, 2005


Davies, G., “A history of money: from ancient times to present day”, *University of Wales Press*, 2002


A Comprehensive Study of Bank Branch Growth Potential and Growth Trends


Podinovski, V., “Production Trade-Offs and Weight Restrictions in Data Envelopment Analysis”, *Journal of the Operational Research Society*, vol.55, pp.1311-1322, 2004


Wilson, P., “FEAR”, *R-Package Software*, 2013


Appendix A: Data Treatment and Variable Selection

To ensure that this research is fully reproducible by the reader, this Appendix will provide a detailed step-by-step list of the data cleaning and treatment process. In addition, the list includes the steps taken to select the appropriate variables for each model developed in the body of this work.

**Data Cleaning:**

Before any methodologies can be applied to a data set, the data must be carefully looked over to identify any missing or suspect data as well as any obvious outlier units. For the purpose of this study, an outlier unit was considered to be a branch that has a unique function (oil and gas, real estate etc.) or a branch that was non-comparable due to unique operating or environmental conditions.

1. Initial data review/cleaning:

   The Turkish banking data set used in this study was obtained from a data analysis firm which had already cleaned the data set of any errors, missing data or outlying units, and thus no preliminary work was required for this data set aside from understanding the nature of the data and the products and branches it represented.

   The Canadian Data set, however, required careful consideration of all variables for each DMU. Any DMU that was missing or had suspect data was removed from the data set. In total, one unit was removed for having missing data, while three were removed for having negative values for strictly positive variables (eg. number of customer, number of products). It is likely that these discrepancies were due to a reporting error. Subsequent to the removal of these units, the maximum, minimum and average values were assessed for each variable to determine whether there were any obvious outlier units in the data set. This lead to the removal of the large downtown Toronto branch which employed over 10,000 employee which was substantially larger than the average of 12 employees reported for the remaining 1035 branches.

   Data cleaning was also required for the Statistics Canada data that was collected for the formulation of the indices. Since there are over 65,000 DAs across Canada reporting...
errors are not uncommon and must be removed or corrected in order to produce meaningful indices. Common errors included missing data and negative population and income values.

**Data Treatment and Variable Selection:**

Once the data is cleaned of any obvious errors and outliers, one can proceed with preliminary model formulation and variable selection. Subsequently, sensitivity analysis techniques should be applied to these preliminary models in order to objectively determine which variables should remain in the model. This process generally results in fewer input and output variables thus improving the discrimination and reducing the dimensionality of the model. Finally, additional outlier removal techniques can be applied to ensure that all outliers have been identified and removed. The following sections outline the exact process applied to each model introduced in this document.

2. Preliminary Input and Output Selection:

Both the Turkish and Canadian data sets contained extensive amounts of product data as well as employment data for each branch but did not contain any cost or transaction data. Consequently, both data sets lent themselves well to use in a production model. Preliminary models were then set up which included all relevant product and employment data and indices where applicable. The preliminary models for the Canadian and Turkish data sets are provided in Tables A1.1 and A1.2, respectively.

**Table A1.1: Preliminary Canadian Bank Variable Selection**

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>Products</td>
</tr>
<tr>
<td>- FTEs</td>
<td>- Amount of Funds</td>
</tr>
<tr>
<td>Indices</td>
<td>- Number of Products</td>
</tr>
<tr>
<td>- Income Index</td>
<td>- Number of Customers</td>
</tr>
<tr>
<td>- Population Index</td>
<td>Competitive Indices</td>
</tr>
<tr>
<td></td>
<td>- 100 KM</td>
</tr>
<tr>
<td></td>
<td>- 50 KM</td>
</tr>
<tr>
<td></td>
<td>- 25 KM</td>
</tr>
<tr>
<td></td>
<td>- 10 KM</td>
</tr>
<tr>
<td></td>
<td>- 5 KM</td>
</tr>
<tr>
<td></td>
<td>- 1 KM</td>
</tr>
</tbody>
</table>
Table A1.2: Preliminary Turkish Bank Variable Selection

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>Number of Customers</td>
</tr>
<tr>
<td></td>
<td>Consumer Loans</td>
</tr>
<tr>
<td></td>
<td>Mortgage Loans</td>
</tr>
<tr>
<td></td>
<td>Demand Deposits</td>
</tr>
<tr>
<td></td>
<td>Term Deposits</td>
</tr>
<tr>
<td></td>
<td>Active Product Usage</td>
</tr>
<tr>
<td></td>
<td>Number of Clients with Active Term Deposits</td>
</tr>
<tr>
<td></td>
<td>Number of Clients with Active Demand Deposits</td>
</tr>
</tbody>
</table>

These preliminary models included all relevant input and output variables that were provided in the data sets.

3. Variable Selection and Sensitivity Analysis

Once the preliminary models were outlined, sensitivity analyses, including correlation analysis, efficiency contribution method and principal component analysis, were performed to remove redundant variables and improve discrimination. Each of these analyses were performed in R and all relevant R-scripts are provided in Appendix B.

Correlation Analysis:

Correlation analysis provides a quick means of investigating the relationships that exist between all inputs and outputs. Using the “corr()” function in R, correlation matrices were obtained for both models. High correlation coefficients (\(|\rho|>0.95\)) between two inputs (outputs) can indicate redundancy, while low correlation coefficients (\(|\rho|<0.50\)) between an input and an output can indicate the lack of a causal relationship. It follows that if either of these circumstances occurred the variables in questions were investigated further. For example, in the Canadian model the correlation coefficients between the product outputs were generally high as one would expect. If a branch has more customers, they will likely have more products being sold and more funds being deposited. Correlation analysis must be used with caution as the removal of variables does not always make sense from a managerial standpoint and will inevitably lead to the removal of information from the model. No variables were removed at this point in the analysis; suspect relationships were noted and further sensitivity analysis was performed to support/refute the findings.

Efficiency Contribution Method:
The efficiency contribution method is a simple way to observe the effects of a variable on the average efficiency score of a model. This method requires running the DEA model with all of the variables, and then re-running the model with one of the variables removed. The average efficiency scores of the two runs are then compared. As employment was the only discretionary input variable in either the Canadian or Turkish model, it was not removed and, thus, only output variables were investigated with this method. Since there exists such a strong relationship between output variables in either model, it was found that the removal of any variables did not have a significant impact on the average efficiency scores. Although there were fluctuations in the efficiency scores, the results of this methodology were not definitive enough to draw any concrete conclusions pertaining to variable removal in either model.

Principal Component Analysis:

Principal component analysis was the main methodology responsible for variable removal in this study. Using the variations of the R-script provided in Appendix B, the efficiency distributions of several models with differing variable selections were compared. If it was found that the removal of a variable has a significant impact on the distribution, the variable would remain in the model. On the other hand, if it were found that the removal of the variable did not have a significant impact on the efficiency distribution, the variable was removed. For the Canadian model, it was found that only the 5KM and 10KM competitive index significantly impacted to efficiency distribution. The 5KM index was arbitrarily chosen for the final model. It was also found that only two product variables were required to maintain the efficiency distribution. However, since dimensionality was not an issue given the large number of DMUs included in the data set, all three variables were included in the final model to provide more detailed target objectives and easily understood results. The final Canadian Bank Model is provided in Table A1.3.
Table A1.3: Final Canadian Bank Model

<table>
<thead>
<tr>
<th>Resources</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>- FTEs</td>
<td>- Amount of Funds</td>
</tr>
<tr>
<td>Indices</td>
<td>- Number of Products</td>
</tr>
<tr>
<td>- Income Index</td>
<td>- Number of Customers</td>
</tr>
<tr>
<td>- Population Index</td>
<td>Competitive Indices</td>
</tr>
<tr>
<td></td>
<td>- 5 KM</td>
</tr>
</tbody>
</table>

PCA analysis applied to the Turkish model resulted in the removal of “Number of Clients with Active Term Deposits”, “Number of Clients with Active Demand Deposits” and “Active Product Usage”. It was also found that either “Number of Customers” or “Demand Deposits” (not both) could be removed, however, for much the same reason as the Canadian model these variables were left in the final model which is presented in Table A1.4.

Table A1.4: Final Turkish Bank Model

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>Number of Customers</td>
</tr>
<tr>
<td></td>
<td>Consumer Loans</td>
</tr>
<tr>
<td></td>
<td>Mortgage Loans</td>
</tr>
<tr>
<td></td>
<td>Demand Deposits</td>
</tr>
<tr>
<td></td>
<td>Term Deposits</td>
</tr>
</tbody>
</table>

4. Additional Outlier Removal

After the final models were selected, additional outlier techniques were performed to ensure that no additional outliers were left in the data set. It should be noted that although the techniques discussed in this section help in the identification of outliers, it does not necessarily mean that all outliers have been removed from the data set. Careful consideration by Bank management would need to be performed to know with 100% certainty that all outliers, including specialty branches, have been removed.

*Stripping the Efficient Frontier:*

To verify that all obvious outliers were removed from both models, the efficient frontier was stripped and the change in the average efficiency score was examined. For both models, it was found, that the changes in average efficiency observed were not significant and thus no other obvious outliers were found to be sitting on the efficient frontier.
Wilson’s Outlier Detection Statistic:

The Wilson’s Outlier Detection Statistics is a useful method of pinpointing possible outliers in a data set. Unfortunately, it requires significant computing power for large data sets and thus could only be run for the small simulated data set and not the Turkish or Canadian bank data sets. The application of this methodology on the test set is described in detail in Section 6.3.2 while the R-script for its application is provided in Appendix B.

DMU Segmentation:

Local analyses can provide new and useful insights about the DMUs in question by separating them into smaller groups of DMUs that share common characteristics. Local analysis was only performed for the Canadian bank model in this study and segmentation was based on the demographic environments and local competition of each branch (i.e. population, income and competitive indices). Using the within group sum of squares the ideal number of clusters was determined and subsequently the k-means clustering technique was applied using the R-script provided in Appendix B. For a detailed description of this process please refer to Section 5.3.3 of this document.

Applying Developed DEA methodologies:

Once data cleaning, variable selection, outlier removal and DMU segmentation were performed, the methodologies developed herein were applied to the data sets using the R-scripts provided in Appendix C and D. The analysis of the results obtained from these methodologies are discussed in detail throughout Chapters 6 and 7.
Appendix B: Data Treatment R-Scripts

Principal Component Analysis (Wilcoxon Rank Sum Test)-

Growth Model Example:

library("ucminf")
library('lpSolveAPI')
library("Benchmarking")
growthdata = read.csv('C:/Users/alex/Desktop/Growth Model.csv')

y= with(growthdata, cbind(pfund2009,puc2009, cust2009))
e <- dea(x,y,RTS='vrs', ORIENTATION='out')
E= 1/eff(e)
emod <- dea(xmod,y,RTS='vrs', ORIENTATION='out')
Emod= 1/eff(emod)
wilcox.test(E,Emod)

Wilson’s Outlier Detection Using Fear Package-

library("FEAR")

#Canadian Banking Data
data1 = read.csv('C:/Users/alex/Desktop/Restricted DEA/Restricted Data.csv')
tx<-with(data, cbind(emp) )   #Define Inputs
ty<-with(data, cbind(cust2010, pfund2010, puc2010))  #Define Outputs
x<-t(tx)
y<-t(ty)
outlier<-ap(X=x,Y=y,NDEL=25)
outlier<-data.frame(outlier$r0, outlier$imat)

#Turkish Banking Data
dataturk = read.csv('C:/Users/alex/Desktop/Restricted DEA/Actual/Turkey.csv')
txturk<-with(dataturk, cbind(staff) )   #Define Inputs
tyturk<-with(dataturk, cbind(clients, cloans, mloans, term, demand))  #Define Outputs
xturk<-t(txturk)
yturk<-t(tyturk)
outturk<-ap(X=xturk,Y=yturk,NDEL=20)
ap.plot(RATIO=outturk$Ratio, NLEN = 20, plot.options = list(pch = 19),
        ylim = NULL, ylab = "log-ratio", xlab = "i", main = "",
        err.check = TRUE)
outturk<-data.frame(outturk$r0, outturk$imat)
write.table(outturk, "C:/Users/alex/Desktop/outturk.txt", sep="/t")
# Test Set

datatest = read.csv('C:/Users/alex/Desktop/Restricted DEA/Test Set/TEST.csv')
txtest<-with(datatest, cbind(emp) )  # Define Inputs
tytest<-with(datatest, cbind(CUST, PUC, PFUND))
xtest<-t(txtest)
ytest<-t(tytest)
outtest<-ap(X=xtest,Y=ytest,NDEL=15)
ap.plot(RATIO=outtest$ratio, NLEN = 15, plot.options = list(pch = 19), ylim = NULL, ylab = "log-ratio", xlab = "i", main = "", err.check = TRUE)
outtest<-data.frame(outtest$r0, outtest$imat)

write.table(outtest, "C:/Users/alex/Desktop/outtest2.txt", sep="\t")

Cluster Analysis

mydata = read.csv('C:/Users/alex/Desktop/Growth Model/Growth Model.csv')
mydata10= with(mydata,cbind(X2010POP, X2010INC, X5km))
mydata11=with(mydata,cbind(X2011POP, X2011INC, X5km))
mydata10 <- scale(mydata10)  # standardize variables
mydata11 <- scale(mydata11)

# Determine number of clusters 2010
wss <- (nrow(mydata10)-1)*sum(apply(mydata10,2,var))
for (i in 2:15) wss[i] <- sum(kmeans(mydata10,centers=i)$withinss)
plot(1:15, wss, type="b", xlab="Number of Clusters",ylab="Within groups sum of squares")

# Determine number of clusters 2011
wss <- (nrow(mydata11)-1)*sum(apply(mydata11,2,var))
for (i in 2:15) wss[i] <- sum(kmeans(mydata11,centers=i)$withinss)
plot(1:15, wss, type="b", xlab="Number of Clusters 2011",ylab="Within groups sum of squares")

# K-Means Cluster Analysis 2010
fit1 <- kmeans(mydata10, 6)  # 6 cluster solution
# get cluster means
aggregate(mydata10,by=list(fit1$cluster),FUN=mean)
# append cluster assignment
mydata10 <- data.frame(mydata10, fit1$cluster)
write.table(mydata10, "C:/Users/alex/Desktop/clusters20106.txt", sep="","")

# K-Means Cluster Analysis 2011
fit2 <- kmeans(mydata11, 6)  # 6 cluster solution
# get cluster means
aggregate(mydata11,by=list(fit2$cluster),FUN=mean)
# append cluster assignment
mydata11 <- data.frame(mydata11, fit2$cluster)
write.table(mydata11, "C:/Users/alex/Desktop/clusters20116.txt", sep="",")
Appendix C: Restricted Growth Model

SECTION 1: R-SCRIPTS

Restricted Model R-Script-
(Note: This script must be modified on a per model basis. The script provided was executed for the Canadian Banking data set.)

library('lpSolveAPI')
data = read.csv('C:/Users/alex/Desktop/Restricted DEA/Canadian/Restricted Data.csv')

Primal<- 0  #0=Dual; 1=Primal
Output<-1  #0=Input Oriented; 1=Output Oriented
VRS <- 1 #0=CRS; 1=VRS
Restriction <- 0 #For Dual: OFF=0; ON=1

growth<-with(data,Growth)
inputs<-with(data, cbind(emp) )   #Define Inputs
outputs<-with(data, cbind(cust2010, pfund2010, puc2010))  #Define Outputs
J<-dim(data)[1]                 #Number of DMUs
I<-dim(inputs)[2]               #Number of Inputs
R<-dim(outputs)[2]              #Number of Outputs

if (Primal==0){
theta <- matrix(999.0,J,1)
lambda <- matrix(999.0,J,J)
}

if (Output==0){
for(d in 1:J)
{
    # first set LP variables specific to DMU you are solving for
    x0 <- as.vector(inputs[d,])
    y0<- as.vector(outputs[d,])
    nvars <- 1+J #number of decison variables
    ncons <- R+I
    c1 <- c(rep(0,R),x0) #1st contraint coeffs
    eqs <- c(rep("=".R),rep("=".I)) #constraint types
    rhs <- c(y0,rep(0,I))
    if(VRS==1)
    {
        ncons <- ncons + 1
        c1 <- c(c1,0)
        eqs <- c(eqs,"=")
        rhs <- c(rhs,1)
    }

    RDEA<-make.lp(ncons,nvars)
}
set.column(RDEA,1,c1)

# now set LP variables for all J DMUs
for(j in 1:J)
{
  yj <- as.vector(outputs[j,])
  xj <- as.vector(inputs[j,])
  cj <- c(yj, xj)
  if(VRS==1)
  {
    cj <- c(cj,1)
  }
  set.column(RDEA,j+1,cj)
}

# RHS, constraint types, obj fun
obj <- c(1,rep(0,J))
set.objfn(RDEA,obj)
set.constr.type(RDEA,eqs)
set.rhs(RDEA,rhs)

# set restrictions on lambda
if (Restriction == 1){
  for(j in 1:J)
  {
    if(growth[j]==0){
      set.bounds(RDEA,upper=0.0,columns=j+1)
    }
  }
}

# col and rownames
slam <- vector()
for(j in 1:J)
{
  slam <- c(slam,paste("lambda",toString(j)))
} 
cnames <- c("theta",slam)
rnames <- vector()
for(k in 1:ncons)
{
  rnames <- c(rnames,paste("cons",toString(k)))
} 
dimnames(RDEA) <- list(rnames,cnames)

# solve
status<-solve(RDEA)

# save variables
vars <- get.variables(RDEA)
theta[d] <- vars[1]
lambda[d,] <- vars[2:(J+1)]

if(Output==1){
  for(d in 1:J)
  {
    # first set LP variables specific to DMU you are solving for
    x0 <- as.vector(inputs[d,])
y0 <- as.vector(outputs[d,])
nvars <- 1+J #number of decision variables
    ncons <- R+I
c1 <- c(rep(0,I),y0) #1st constraint coeffs
eqs <- c(rep("\leq",I),rep("\leq",R)) #constraint types
rhs <- c(x0,rep(0,R))
  }
if(VRS==1)
  {
    ncons <- ncons + 1
c1 <- c(c1,0)
eqs <- c(eqs,"=")
rhs <- c(rhs,1)
  }
RDEA<-make.lp(ncons,nvars)
set.column(RDEA,1,c1)

# now set LP variables for all J DMUs
for(j in 1:J)
{
  yj <- as.vector(outputs[j,])
xj <- as.vector(inputs[j,])
cj <- c(xj,-yj)
if(VRS==1)
  {
    cj <- c(cj,1)
  }
  set.column(RDEA,j+1,cj)
}

# RHS, constraint types, obj fun
obj <- c(1,rep(0,J))
set.objfn(RDEA,obj)
set.constr.type(RDEA,eqs)
set.rhs(RDEA,rhs)

# set restrictions on lambda
if (Restriction == 1){

for(j in 1:J)
{
    if(growth[j]==0){
        set.bounds(RDEA,upper=0.0,columns=j+1)
    }
}

# col and rownames
slam <-vector()
for(j in 1:J)
{
    slam <- c(slam,paste("lambda",toString(j)))
}
cnames <- c("theta",slam)
rnames <- vector()
for(k in 1:ncons)
{
    rnames <- c(rnames,paste("cons",toString(k)))
}
dimnames(RDEA) <- list(rnames,cnames)

# solve
lp.control(RDEA, sense='max')
status<-solve(RDEA)

# save varibles
vars <- get.variables(RDEA)
theta[d] <- 1/(vars[1])
lambda[d,] <- vars[2:(J+1)]

}
}
}

if (Primal==1){

xweight <- matrix(999.0,J,I)
yweight <- matrix(999.0,J,R)
w0 <- matrix(999.0,J,1)

if (Output==0){
    for(d in 1:J)
    {
        # first set LP variables specific to DMU you are solving for
        nvars <- R+I #number of decison variables
        ncons <- J+1 #number of constraints

        if(VRS==1)
nvars <- nvars + 1

idxs <- 1:J
if(growth[d]==0)
{
  ncons <- J
  idxs <- c(1:(d-1),(d+1):J)
}

RDEA<-make.lp(ncons,nvars)

c<-0
for(r in 1:R)
{
  c<-c+1
  yr <- outputs[,r]
  cr <- c(yr[1:J],0)
  set.column(RDEA,c,cr)
}
for(i in 1:I)
{
  c<-c+1
  xi <- inputs[,i]
  ci <- c(-xi[1:J],xi[d])
  set.column(RDEA,c,ci)
}

if(VRS==1)
{
  c<-c+1
  clast <- c(rep(-1,length(idxs)),0)
  set.column(RDEA,c,clast)
}

eqs <- c(rep("<=",length(idxs)),"=") #constraint types
rhs <- c(rep(0,length(idxs)),1)

# RHS, constraint types, obj fun
yd <- outputs[,d]
obj <- c(yd,rep(0,I))
if(VRS==1)
{
  obj <- c(obj,-1)
  set.bounds(RDEA,lower=-Inf,columns=(I+R+1))
}

set.objfn(RDEA,obj)
set.constr.type(RDEA,eqs)
set.rhs(RDEA,rhs)
# col and rownames
cnames <- vector()
for(r in 1:R)
{
    cnames <- c(cnames,paste("mu",toString(r)))
}
for(i in 1:I)
{
    cnames <- c(cnames,paste("nu",toString(i)))
}
if(VRS==1)
{
    cnames <- c(cnames,"w0")
}

rnames <- vector()
for(idx in idxs)
{
    rnames <- c(rnames,paste("dmu",toString(idx)))
}
rnames <- c(rnames,"sumx0")
dimnames(RDEA) <- list(rnames,cnames)

# solve
lp.control(RDEA, sense='max')
status<-solve(RDEA)

# save variables
vars <- get.variables(RDEA)
yweight[d,] <- vars[1:R]
xweight[d,] <- vars[(R+1):(R+I)]
if(VRS==1)
{
    w0[d,] <- vars[(R+I+1)]
}

if (Output==1){
    for(d in 1:J)
    {
        # first set LP variables specific to DMU you are solving for
        nvars <- R+I #number of decision variables
        ncons <- J+1 #number of constraints

        if(VRS==1)
        {
            nvars <- nvars + 1
        }
    }
idxs <- 1:J
if(growth[d]==0)
{
    ncons <- J
    idxs <- c(1:(d-1),(d+1):J)
}

RDEA<-make.lp(ncons,nvars)

<=0
for(r in 1:R)
{
    <=c+1
    yr <- outputs[,r]
    cr <- c(yr[idxs],yr[d])
    set.column(RDEA,c,cr)
}
for(i in 1:I)
{
    <=c+1
    xi <- inputs[,i]
    ci <- c(-xi[idxs],0)
    set.column(RDEA,c,ci)
}

if(VRS==1)
{
    <=c+1
    clast <- c(rep(1,length(idxs)),0)
    set.column(RDEA,c,clast)
}

# RHS, constraint types, obj fun
eqs <- c(rep("="+length(idxs)),')=',1) #constraint types
rhs <- c(rep(0,length(idxs)),1)

xd <- inputs[d]
obj <- c(rep(0,R),xd)
if(VRS==1)
{
    obj <- c(obj,-1)
    set.bounds(RDEA,lower= -Inf,columns=(I+R+1))
}

set.objfn(RDEA,obj)
set.constr.type(RDEA,eqs)
set.rhs(RDEA,rhs)

# col and rownames
cnames <- vector()
for(r in 1:R)
{
  cnames <- c(cnames,paste("mu",toString(r)))
}
for(i in 1:I)
{
  cnames <- c(cnames,paste("nu",toString(i)))
}
if(VRS==1)
{
  cnames <- c(cnames,"w0")
}

rnames <- vector()
for(id in idxs)
{
  rnames <- c(rnames,paste("dmu",toString(id)))
}  
rnames <- c(rnames,"sumx0")  
dimnames(RDEA) <- list(rnames,cnames)

# solve
lp.control(RDEA, sense='min')
status<-solve(RDEA)

# save varibles
vars <- get.variables(RDEA)
yweight[d,] <- vars[1:R]
xweight[d,] <- vars[(R+1):(R+I)]
if(VRS==1)
{
  w0[d] <- vars[(R+I+1)]
}

write.table(theta,"C:/Users/alex/Desktop/theta.txt",sep="\t")
write.table(lambda,"C:/Users/alex/Desktop/lambda.txt",sep="\t")

EXFA Script-

library('lpSolveAPI')
library('Benchmarking')
source('C:/Users/Alex/Desktop/slacks.R')

#CCR example from Olesen and Petersen
#inputs <- matrix(rep(1.0,10))
#outputs <- cbind(
#    matrix(c(5,10,50,80,100,115,125,10,42,85),
# matrix(c(10,5,50,80,100,125,115,15,42,85)),
# matrix(c(120,120,110,100,90,1,1,55,45,95)))
# VRS <- 0

# BCC example from Olesen and Petersen
# inputs <- cbind(
#   matrix(c(5,20,30,40,50,130,110,20,100,40)),
#   matrix(c(20,5,30,40,50,110,130,100,100,40)))
# outputs <- matrix(c(10,10,65,80,90,120,130,15,80,70))
# VRS <- 1

# textbook example
# inputs <- cbind(
#   matrix(c(10,15,20,25,12)),
#   matrix(c(20,15,30,15,9)))
# outputs <- cbind(
#   matrix(c(70,100,80,100,90)),
#   matrix(c(6,3,5,2,8)))
# VRS <- 0

# Canadian Data Set
# data = read.csv('C:/Users/Alex/Desktop/Restricted DEA/Canadian/Restricted Data.csv')
growth <- with(data, Growth)
inputs <- with(data, cbind(emp))  # Define Inputs
outputs <- with(data, cbind(cust2010, pfund2010, puc2010, POP, INC, X5km))  # Define Outputs
VRS <- 0  # 0=CRS; 1=VRS

# Turkish Data Set
# data = read.csv('C:/Users/Alex/Desktop/Restricted DEA/Turkish/Turkey.csv')
growth <- with(data, Growth)
inputs <- with(data, cbind(staff))  # Define Inputs
outputs <- with(data, cbind(clients, cloans, mloans, term, demand))  # Define Outputs
VRS <- 0  # 0=CRS; 1=VRS

# set dimensions of optimization problem
J <- dim(inputs)[1]  # number of DMUS
I <- dim(inputs)[2]  # number of inputs
R <- dim(outputs)[2]  # number of outputs

# just benchmarking package to get base CCR or BCC results
first <- 0
rts <- "crs"
if(VRS == 1) {
  rts <- "vrs"
}
first <- dea(inputs, outputs, ORIENTATION="in", RTS=rts, SLACK=TRUE)

# store baseline results
theta_first <- eff(first)
s <- slacks(inputs,outputs,theta_first,VRS)
strong_eff <- (theta_first==1 & s$sum == 0)

# note: benchmarking package calculates slacks incorrectly! use function
# "slacks" in slacks.R script instead!
E <- which(strong_eff==TRUE)
size_E <- length(E)

# define matrices to hold results of EXFA model
xweight <- matrix(999.0,J,I)
yweight <- matrix(999.0,J,R)
#sj <- matrix(999.0,J,size_E)
binaries <- matrix(999.0,J,size_E)
w0 <- matrix(999.0,J,1)
theta <- rep(999.0,J)

# EXFA model
M<-1000000000000
for(d in 1:J)
{

  nvars <- R+I+size_E*2 # number of decision variables: mu, nu, slacks,binaries
  ncons <- size_E*2+2 # number of constraints

  # if variable returns to scale, add additional variable w0
  if(VRS==1)
  {
    nvars <- nvars + 1
  }

  idxs <- E
  LP<-make.lp(ncons,nvars)

  # loop through columns of constraint set...
  c<-0
  constraint_mat <- matrix(,ncons)

  # ...first R columns represent mu (output multipliers)
  for(r in 1:R)
  {
    c<-c+1
    yr <- outputs[,r]
    cr <- c(yr[idxs],0,rep(0,size_E),0)
    set.column(LP,c,cr)
    constraint_mat <- cbind(constraint_mat,matrix(cr))
  }

  # ...second I columns represent nu (input multipliers)
  for(i in 1:I)
  {
    c<-c+1
  }
xi <- inputs[,i]
ci <- c(-xi[idxs],xi[d],rep(0,size_E),0)
set.column(LP,c,ci)
constraint_mat <- cbind(constraint_mat,matrix(ci))
}
# ...next J columns represent s (slacks)
for(j in idxs)
{
  c<-c+1
  tmp <- rep(0,size_E)
  for(k in 1:size_E)
  {
    if(idxs[k]==j) {tmp[k]=1}
  }
  cj <- c(tmp,0,tmp,0)
  set.column(LP,c,cj)
  constraint_mat <- cbind(constraint_mat,matrix(cj))
}
# ...last J columns represent binaries
for(j in idxs)
{
  c<-c+1
  tmp <- rep(0,size_E)
  for(k in 1:size_E)
  {
    if(idxs[k]==j) {tmp[k]=-M}
  }
  cj <- c(rep(0,size_E),0,tmp,1)
  set.column(LP,c,cj)
  constraint_mat <- cbind(constraint_mat,matrix(cj))
}
# ...if we have VRS, one more column for w0
if(VRS==1)
{
  c<-c+1
  clast <- c(rep(1,size_E),0,rep(0,size_E),0)
  set.column(LP,c,clast)
  constraint_mat <- cbind(constraint_mat,matrix(clast))
}
# drop first col from constraint_mat
constraint_mat <- constraint_mat[,,-1]

# all constraints are equalities in slacks model
eqs <- c(rep("=",size_E),"="
# RHS of constraints are zero except for last one
const <- size_E-(R+I-1)
if(VRS==1)
const <- size_E-(R+I) # IS THIS RIGHT?? SEE PAGE 29... edit: seems so
}
rhs <- c(rep(0,size_E),1,rep(0,size_E),const)

# objective function: maximize sum_{r in R} (mu_r * y_0r)
yd <- outputs[d,]
obj <- c(yd,rep(0,I),rep(0,size_E),rep(0,size_E))
if(VRS==1)
{
  obj <- c(obj,1)
  set.bounds(LP,lower= -Inf,columns=(I+R+2*size_E+1))
}

# set bounds on slacks
#set.bounds(LP,lower=rep(0.0,size_E), columns=(R+I+1):(R+I+size_E))

# restrict binaries
set.type(LP,columns=(R+I+size_E+1):(R+I+size_E*2),type="binary")

# set objective function, constraint types, etc.
set.objfn(LP,obj)
set.constr.type(LP,eqs)
set.rhs(LP,rhs)

# set column names (names of variables)
cnames <- vector()
for(r in 1:R)
{
  cnames <- c(cnames,paste("mu",toString(r)))
}
for(i in 1:I)
{
  cnames <- c(cnames,paste("nu",toString(i)))
}
for(j in E)
{
  cnames <- c(cnames,paste("s",toString(j)))
}
for(j in E)
{
  cnames <- c(cnames,paste("b",toString(j)))
}
if(VRS==1)
{
  cnames <- c(cnames,"w0")
}

# set row names (constraints)
rnames <- vector()
for(idx in idxs) {
  rnames <- c(rnames, paste("dmusum", toString(idx)))
}
rnames <- c(rnames, "sumx0")
for(idx in idxs) {
  rnames <- c(rnames, paste("b", toString(idx)))
}
rnames <- c(rnames, "bsum")
dimnames(LP) <- list(rnames, cnames)

# solve LP!
lp.control(LP, sense='max')
status<-solve(LP)

# save variables
vars <- get.variables(LP)
yweight[d,] <- vars[1:R]
xweight[d,] <- vars[(R+1):(R+I)]
binaries[d,] <- vars[(R+I+size_E+1):(R+I+size_E*2)]
if(VRS==1) {
  w0[d,] <- vars[(R+I+size_E*2+1)]
}

# calculate efficiency
theta[d] <- t(yweight[d,]) %*% yd
if(VRS==1) {
  theta[d] <- theta[d]+w0[d]
}
effs <- cbind(matrix(theta_first), matrix(theta), growth)

write.table(effs, "C:/Users/alex/Desktop/EXFA-DEA.txt", sep="\t")

‘Slacks’ Function Called by EXFA Script-

slacks <- function(inputs, outputs, theta, VRS) {
  sx <- matrix(999.0, J, I)
  sy <- matrix(999.0, J, R)
  sum <- rep(999.0, J)
  for(d in 1:J) {
    x0 <- as.vector(inputs[d,])
    y0<- as.vector(outputs[d,])
    theta0<-theta[d]

    nvars <- J+I+R
    ncons <- I+R

eqs <- c(rep("="),rep("="))
rhs <- c(theta0*x0,y0)
obj <- c(rep(0,J),rep(1,I),rep(1,R))

if(VRS==1) {
  ncons <- ncons + 1
  eqs <- c(eqs,"=")
  rhs <- c(rhs,1)
}

RDEA< make.lp(ncons,nvars)
set.objfn(RDEA,obj)
set.constr.type(RDEA,eqs)
set.rhs(RDEA,rhs)

c <- 0
for(j in 1:J) {
  c <- c+1
  yj <- as.vector(outputs[j,])
  xj <- as.vector(inputs[j,])
  cj <- c(xj,yj)
  if(VRS==1) { cj <- c(cj,1) }
  set.column(RDEA,c,cj)
}

for(i in 1:I) {
  c <- c+1
  tmp <- rep(0,I)
  tmp[i] <- -1
  ci <- c(tmp,rep(0,R))
  if(VRS==1) { ci <- c(ci,0) }
  set.column(RDEA,c,ci)
}

for(r in 1:R) {
  c <- c+1
  tmp <- rep(0,R)
  tmp[r] <- -1
  cr <- c(rep(0,I),tmp)
  if(VRS==1) { cr <- c(cr,0) }
  set.column(RDEA,c,cr)
}

status< solve(RDEA)
vars <- get.variables(RDEA)
sx[d,] <- vars[(J+1):(J+I)]
sy[d,] <- vars[(J+I+1):(J+I+R)]
sum[d] <- sum(sx[d,])+sum(sy[d,])

return(list(sx=sx,sy=sy,sum=sum))
}
### SECTION 2: MODEL RESULTS

#### Demonstrative Application Results

<table>
<thead>
<tr>
<th>DMU</th>
<th>Growth</th>
<th>CRS</th>
<th>RCRS</th>
<th>VRS</th>
<th>RVRS</th>
<th>( \Sigma \Lambda )</th>
<th>RTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.687282</td>
<td>0.694328</td>
<td>0.951319</td>
<td>0.965003</td>
<td>2.749128</td>
<td>DRS</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.715438</td>
<td>0.71924</td>
<td>0.75</td>
<td>0.783477</td>
<td>0.903572</td>
<td>IRS</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.374614</td>
<td>0.380561</td>
<td>1</td>
<td>1</td>
<td>0.374614</td>
<td>IRS</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0.75427</td>
<td>0.803691</td>
<td>0.980576</td>
<td>0.991943</td>
<td>2.262811</td>
<td>DRS</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.557764</td>
<td>0.56253</td>
<td>0.6</td>
<td>0.6</td>
<td>0.348602</td>
<td>IRS</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0.856654</td>
<td>0.857518</td>
<td>0.983112</td>
<td>0</td>
<td>2.645065</td>
<td>DRS</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.099933</td>
<td>0.103833</td>
<td>0.75</td>
<td>0.75</td>
<td>0.099933</td>
<td>IRS</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.397506</td>
<td>0.463374</td>
<td>0.602459</td>
<td>0.602459</td>
<td>1.325018</td>
<td>DRS</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>1</td>
<td>1.008544</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>CRS</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1.165705</td>
<td>1</td>
<td>1.253071</td>
<td>1</td>
<td>CRS</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0.351579</td>
<td>0.39458</td>
<td>0.351944</td>
<td>0.400792</td>
<td>0.896488</td>
<td>IRS</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0.154729</td>
<td>0.156194</td>
<td>0.75</td>
<td>0.75</td>
<td>0.206306</td>
<td>IRS</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0.314013</td>
<td>0.366047</td>
<td>0.584153</td>
<td>0.584153</td>
<td>1.570066</td>
<td>DRS</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0.566389</td>
<td>0.566389</td>
<td>0.792383</td>
<td>0.566389</td>
<td>2.262811</td>
<td>DRS</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0.60213</td>
<td>0.67927</td>
<td>0.636191</td>
<td>0.74914</td>
<td>1.285277</td>
<td>DRS</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>0.549567</td>
<td>0.564519</td>
<td>1</td>
<td>1</td>
<td>2.747833</td>
<td>DRS</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0.324322</td>
<td>0.36408</td>
<td>0.328862</td>
<td>0.382801</td>
<td>1.081073</td>
<td>DRS</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>CRS</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>0.881579</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.644737</td>
<td>DRS</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0.495932</td>
<td>0.498959</td>
<td>1</td>
<td>1</td>
<td>0.398721</td>
<td>IRS</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0.2431</td>
<td>0.277101</td>
<td>0.75</td>
<td>0.75</td>
<td>0.324133</td>
<td>IRS</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0.451754</td>
<td>0.512438</td>
<td>0.511603</td>
<td>0.677632</td>
<td>0.677632</td>
<td>IRS</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>0.333175</td>
<td>0.338122</td>
<td>0.432894</td>
<td>0.43384</td>
<td>1.468197</td>
<td>DRS</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>CRS</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>0.607928</td>
<td>0.622712</td>
<td>0.75</td>
<td>0.775799</td>
<td>0.810571</td>
<td>IRS</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>0.517117</td>
<td>0.528136</td>
<td>0.61404</td>
<td>0.61404</td>
<td>1.292793</td>
<td>DRS</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0.21394</td>
<td>0.249391</td>
<td>0.6</td>
<td>0.6</td>
<td>0.356566</td>
<td>IRS</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>0.313442</td>
<td>0.325674</td>
<td>0.6</td>
<td>0.6</td>
<td>0.391802</td>
<td>IRS</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>0.871145</td>
<td>0.872103</td>
<td>1</td>
<td>0</td>
<td>2.6704</td>
<td>DRS</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.543176</td>
<td>0.557797</td>
<td>0.6</td>
<td>0.65528</td>
<td>0.894533</td>
<td>IRS</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>0.334721</td>
<td>0.343244</td>
<td>0.351401</td>
<td>0.390255</td>
<td>1.56203</td>
<td>DRS</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>0.544546</td>
<td>0.585695</td>
<td>1</td>
<td>1</td>
<td>0.3081</td>
<td>IRS</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>0.515351</td>
<td>0.58506</td>
<td>0.540267</td>
<td>0.586949</td>
<td>1.288377</td>
<td>DRS</td>
</tr>
<tr>
<td>34</td>
<td>1</td>
<td>0.290106</td>
<td>0.296778</td>
<td>0.32914</td>
<td>0.353883</td>
<td>1.740634</td>
<td>DRS</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>0.613573</td>
<td>0.695994</td>
<td>1</td>
<td>0</td>
<td>2.914474</td>
<td>DRS</td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>0.415021</td>
<td>0.416099</td>
<td>0.420676</td>
<td>0.463956</td>
<td>1.278921</td>
<td>DRS</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.141988</td>
<td>1</td>
<td>1.156055</td>
<td>1</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>0.767724</td>
<td>0.804095</td>
<td>0.794355</td>
<td>1</td>
<td>1.247837</td>
<td>DRS</td>
</tr>
<tr>
<td>39</td>
<td>1</td>
<td>0.50679</td>
<td>0.515818</td>
<td>0.903396</td>
<td>0.903396</td>
<td>2.146448</td>
<td>DRS</td>
</tr>
</tbody>
</table>
Appendix D: Period Growth Model

SECTION 1: R-SCRIPTS

Global Malmquist-

library("ucminf")
library("lpSolveAPI")
library("Benchmarking")

# dataset should be structured as follows:
# column 1: identifier (id)
# column 2: time period (year)
# columns 3-(2+I) inputs (there are I of them)
# columns (3+I)-(2+R+I): outputs (there are R of them)
# example:
# sequence | year | x1 | x2 | y1 | y2 | y3

cat("Loading data...\n")
dat = read.csv('C:/Users/alex/Desktop/Growth GM.csv')

I<6 # ninputs
R<3 # noutputs

orient <- "out"
scale <- "crs"

# unique list of ids and years
ids <- unique(dat[,"id"])
years <- sort(unique(dat[,"year"]))
'years <- years[1:(length(years)-1)]'

# output-input cols
x.idx <- 3:(2+I)
y.idx <- (3+I):(2+R+I)

# empty columns to store distances
dat2 <- dat
dat2["D_t1_t1"] <- NA
dat2["D_t_t"] <- NA
dat2["D_G_t1"] <- NA
dat2["D_G_t"] <- NA

# first t to t ad G to t
cat("Running DEA to calculate D_t_t and D_G_t...
")
for (t in years)
{

cat("\n Year ",t,"...
")
# t to t
tmp <- dat[dat["year"]==t,]
basex <- tmp[, x.idx]
basey <- tmp[, y.idx]
frontierx <- tmp[, x.idx]
frontiery <- tmp[, y.idx]
cat("\n\n")
}
res <- dea(X=baseX,Y=baseY,ORIENTATION=orient, RTS=scale, FAST=TRUE)
res2 <- res
if(orient=="in") {res2<-1/res2}
dat2[dat2[,"year"]==t,"D_G_t"] <- res2

# G to t
frontiex <- dat[,x.idx]
frontiery <- dat[,y.idx]
cat("t" D_G_t"
res <- dea(X=baseX,Y=baseY,XREF=frontiex,YREF=frontiery,
ORIENTATION=orient, RTS=scale, FAST=TRUE)
res2 <- res
if(orient=="in") {res2<-1/res2}
dat2[dat2[,"year"]==t,"D_G_t"] <- res2
}

# now loop through all ids and years (except the first one),
# and shift t+1 variables back by one year


for (id in ids)
{
    for (t1 in years[-1])
    {
        t1.idx <- (dat["id"]==id & dat["year"]==t1)
        t.idx <- (dat["id"]==id & dat["year"]==(t1-1))
        dat2[t.idx,"D_t1_t1"] <- dat2[t1.idx,"D_t_t"]
        dat2[t.idx,"D_G_t1"] <- dat2[t1.idx,"D_G_t"]
    }
}

# calculate EC and FS
dat2["EC"] <- dat2["D_t1_t1"] / dat2["D_t_t"]
dat2["FS"] <- {dat2["D_G_t1"] * dat2["D_t_t"]} / {dat2["D_t1_t1"] + dat2["D_G_t"]}
dat2["malm"] <- dat2["EC"] * dat2["FS"]

Re-Shaping Data Set for Rolling Window Analysis-

myData = read.csv('C:/Users/alex/Desktop/Growth Model-Window.csv')

# identify the variables we are interested in creating leads of and reshaping
vars<-c("INC","POP","ncust","puc","pfund")

# here are the years in which we have observations for the variables

# create the lead variables for all years except the last one since we don't have data beyond
# the last year
for(var in vars)
{
    for(year in years[0:(length(years)-1)])
    {
        sourcevar <- paste(var,toString(year+1),sep=".")
        newvar <- paste(var,"next",sep="_")
        newvar <- paste(newvar,toString(year),sep=".")
        myData[newvar] <- myData[sourcevar]
    }
}
for(var in vars)
{
    newvar <- paste(paste(var,"next",sep="_"),toString(years[length(years)]),sep=".")
    myData[newvar] <- NaN
}

# create the list of variables we want to reshape
varying_vars <- c()
for(var in vars)
{
    for(year in years)
    {
        varying_vars <- c(varying_vars,paste(var,year,sep="."))
        varying_vars <- c(varying_vars,paste(paste(var,"next",sep="_"),year,sep="."))
    }
}

# reshape!
newData <- reshape(myData, idvar='Sequence', varying = varying_vars, sep='.', direction='long')
write.table(newData, "C:/Users/alex/Desktop/newdata.txt", sep=" ", row.names=FALSE)

Rolling Window Analysis
library("ucminf")
library('lpSolveAPI')
library('Benchmarking')

winddata = read.csv('C:/Users/alex/Desktop/newdata.csv')

'2008-2010'
subset1 <- subset(winddata, time<=2010 & time>=2008)
x1 = with(subset1, cbind(X.emp, pfund,puc, ncust, POP, INC) )
y1 = with(subset1, cbind(pfund_next,puc_next, ncust_next))
e1 <- dea(x1,y1,RTS='vrs', ORIENTATION='out')
E1= 1/eff(e1)

'2009-2011'
subset2 <- subset(winddata, time<=2011 & time>=2009)
x2 = with(subset2, cbind(X.emp, pfund,puc, ncust, POP, INC) )
y2 = with(subset2, cbind(pfund_next,puc_next, ncust_next))
e2 <- dea(x2,y2,RTS='vrs', ORIENTATION='out')
E2= 1/eff(e2)

'2010-2012'
subset3 <- subset(winddata, time<=2012 & time>=2010)
x3 = with(subset3, cbind(X.emp, pfund,puc, ncust, POP, INC) )
y3 = with(subset3, cbind(pfund_next,puc_next, ncust_next))
e3 <- dea(x3,y1,RTS='vrs', ORIENTATION='out')
E3= 1/eff(e3)
SECTION 2: MODEL RESULTS
Efficiency Distributions for Complete Data Set - 2008-2009 Model

2009-2010 Model

2010-2011 Model
2011-2012 Model

![Histogram for 2011-2012 Model](image1)

2012-2013 Model

![Histogram for 2012-2013 Model](image2)
Appendix E: Preliminary Work

The following paper documents the preliminary work performed in relation to the research documented in this thesis: