

International Series in  
Operations Research & Management Science

Joseph C. Paradi  
H. David Sherman  
Fai Keung Tam

# Data Envelopment Analysis in the Financial Services Industry

A guide for practitioners and analysts  
working in Operations Research  
using DEA



 Springer

# **International Series in Operations Research & Management Science**

Volume 266

## **Series Editor**

Camille C. Price

Stephen F. Austin State University, TX, USA

## **Associate Series Editor**

Joe Zhu

Worcester Polytechnic Institute, MA, USA

## **Founding Series Editor**

Frederick S. Hillier

Stanford University, CA, USA

More information about this series at <http://www.springer.com/series/6161>

Joseph C. Paradi • H. David Sherman  
Fai Keung Tam

# Data Envelopment Analysis in the Financial Services Industry

A Guide for Practitioners and Analysts  
Working in Operations Research Using DEA

Joseph C. Paradi  
University of Toronto  
Toronto, ON, Canada

H. David Sherman  
Northeastern University  
Boston, MA, USA

Fai Keung Tam  
University of Toronto  
Toronto, ON, Canada

ISSN 0884-8289 ISSN 2214-7934 (electronic)  
International Series in Operations Research & Management Science  
ISBN 978-3-319-69723-9 ISBN 978-3-319-69725-3 (eBook)  
<https://doi.org/10.1007/978-3-319-69725-3>

Library of Congress Control Number: 2017955958

© Springer International Publishing AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To our significant others: Monika, Linda,  
and Bernice; David's supportive daughters  
Amanda and Caroline; and Joseph's sons  
Joseph and David and grandchildren  
Andrew, Laura, and Sophie*

# Foreword

I am extremely pleased that the authors have written this book and honored that they have invited me to develop the foreword.

I have known Joe Paradi and David Sherman for a number of years and have followed their pioneering research from the beginning. Joe was the first to utilize extensive data visualization to communicate DEA results to managers, while David wrote the first introductory monograph explaining DEA for the service sector. Both have extensive consulting experience and managerial expertise which produce a unique and valuable perspective. Over the years, they have developed separate impressive research agendas. I am very pleased that they have joined forces with Fai Keung Tam to produce this book. We have had many discussions of the critical need for such a book collecting together and showcasing studies of managerial importance. The result should help the reader better appreciate the power of DEA as a novel approach for organizing and analyzing data to produce valuable insight.

As mentioned in their introduction, there has been a host of DEA-related articles produced in the past 40 years. The DEA bibliography that I maintain now contains around 15,000 books, dissertations, and articles published since 1978. Unfortunately, the majority of these articles are not particularly useful. Many are a simple study of a specific industry in a single country at one point in time for which the results simply state the relative efficiency scores for a list of DMUs. Such articles are not valuable in that they are a simple ordering of units and do not provide helpful insight for managers such as trends, comparisons across regions, organizational subgroups or ownership types, multinational comparisons, etc. In short, the explanatory power is small frequently due to the shortage of temporal data, failure to perform a thorough analysis across multiple models, model extensions, and various subsets of the data and/or shortcomings of the experimental design.

This book seeks to address this problem by showcasing articles from the financial services area that describe innovative approaches and novel applications that provide insight and uncover transferable best practices. Of course the models, approaches, and advice while stated in the context of financial services are easily applicable to other industry sectors.

My hope is that DEA researchers will familiarize themselves with these compelling applications and approaches and heed the authors' guidance and advice. Hopefully this will result in a significant increase in the number of useful DEA articles for which rigorous analysis produces valuable insight and directly impacts managerial practice. Such an advancement will enhance the field and more fully realize the potential of the DEA methodology.

Lawrence M. Seiford  
Department of Industrial and Operations  
Engineering  
The University of Michigan  
Ann Arbor, MI, USA

# Acknowledgments

Once a person writes a book, s/he learns the hard lesson that it takes several times more effort than anticipated, longer than planned and the completion almost becomes an obsession. Of course, aside from the authors of the book, a lot of others make contributions, some more, some less, but all are essential to success.

First, we would like to thank the owners of the copyrighted materials they so graciously allowed us to use and include in this book. All good contributions to science are based on the work that has been done by many others in the far as well as in the recent past.

Aside from the use of copyrighted materials from external sources, we made use of a substantial amount of research results and work completed or being worked on by our own students. These outstanding young women and men form the foundation of the future in not only DEA but all aspects of our society. We appreciate their enthusiasm in helping us with this book. They deserve much of the credit for the ideas, development, and progress in the application of DEA to real-life problems. Here they are and our postdoctoral fellows:

Burçu Anadol	Parisa H. Ardehali	Maryam Badrizadeh
Barak Edelstein	Allison Hewlitt	Angela Tran Kingyens
Alex E. LaPlante	Denise McEachern	Elizabeth Min
Peter Pille	Stephen Rouatt	Paul C. Simak
Shabnam Sorkhi	Taraneh Sowlati	Niloofar Tochaie
Sandra A. Vela	D’Andre Wilson	Tracy Yang
Zijiang Yang		

Postdoctoral fellows: Mette Asmild, Dan Rosen, Claire Schaffnit, Xiaopeng Yang, and Haiyan Zhu

We also thank all who suggested ways to do things, provided examples of how to view real-world problems, and added the “reality” factor to the work we reported on. We thank dozens of professionals who collaborated with us in the work, without



whom the underlying research and therefore this book could not have been completed. Among these stands out David Paradi who is a master at using PowerPoint and has contributed his knowledge and enthusiasm to the production of figures we present here, and many other technical issues.

A special thank you is due to Professor Joe Zhu who suggested to us that a book like this was needed and then answered all our questions. He is one of today's most respected authors and authorities on DEA. Very much is owed to our late friend, Prof. W.W. (Bill) Cooper, who was one of the creators of DEA and was first to introduce us to the boundless problem-solving capabilities of this excellent tool.

And last, but not least, we thank our better halves Monika Paradi, Linda Sherman, and Bernice Cheng for their patience and even encouragements while they were neglected during the creation of this book.

# Contents

## Part I Data Envelopment Analysis, in Brief with Little Math!

<b>1</b>	<b>DEA Models Overview</b>	<b>3</b>
	Basic DEA Models	4
	Introduction	4
	Model Types	5
	Radial Models: CCR	5
	Radial Models: BCC	9
	Additive Models	11
	SBM Model	12
	Practical Extensions	13
	Input, Output, and Data Issues	16
	Inputs and Outputs	16
	The DEA “Family Tree”: Evolution of Applications and Methodology	17
	Summaries of DEA Research and Publications	18
	Methodological Developments	20
	Application Developments	23
	First Use of DEA in Banking by Topic: DEA Banking Timeline	26
	Appendix: Chapter 1 (Sherman and Zhu 2006)	30
	How DEA Works and How to Interpret the Results	30
	The Mathematical Formulation of DEA	34
	References	37
<b>2</b>	<b>Survey of the Banking Literature</b>	<b>41</b>
	Introduction	41
	Literature Pertinent to This Work	41
	References	47

<b>3</b>	<b>Survey of Other Financial Services Literature</b>	51
	Introduction	51
	Thriffs and Similar Institutions	51
	Insurance	53
	Investment Funds (Mutual Funds, Hedge Funds and Pension Funds)	56
	Mutual Funds	56
	Hedge Funds	58
	Pension Funds	59
	Stock Selection	61
	Concluding Remarks	63
	References	64

## **Part II DEA in Banking**

<b>4</b>	<b>Banking Corporation Studies: In-Country Studies</b>	71
	Introduction	71
	Case 1: Indonesia	71
	Case 2: India	73
	Different Points of View Result in Different Outcomes	75
	References	77
<b>5</b>	<b>Banking Corporation Studies: Multinational Studies</b>	79
	Introduction	79
	Cross-Country Bank Branch Comparisons	79
	References	86
<b>6</b>	<b>Bank Branch Productivity Applications: Basic Applications – Efficiency Measurement</b>	87
	Introduction	87
	References	99
<b>7</b>	<b>Bank Branch Productivity Applications: Managing Bank Productivity</b>	101
	Introduction	101
	Applying DEA to Growth Bank	102
	Specifying Resource Inputs and Service Outputs	103
	DEA Branch Productivity Results	105
	Conclusions	111
	References	111
<b>8</b>	<b>Bank Branch Productivity Applications: Focused Applications to Improve Performance</b>	113
	Introduction	113
	Improvement Targets for Efficient DMUs	117
	References	127

<b>9 Bank Branch Productivity Applications: Strategic Branch Management Issues Addressed with DEA . . . . .</b>	<b>129</b>
Introduction . . . . .	129
Concluding Remarks . . . . .	141
References . . . . .	142
<b>10 Bank Branch Operational Studies Using DEA . . . . .</b>	<b>145</b>
Introduction . . . . .	145
Mergers and Acquisition: Potential Use of DEA to Monitor and Manage the Process . . . . .	145
Product Efficiency and Business Growth . . . . .	152
Conclusions . . . . .	157
References . . . . .	157
<b>11 Bank Branch Benchmarking with Quality as a Component . . . . .</b>	<b>159</b>
Introduction . . . . .	159
Topic 1: Incorporating Quality Variables into a DEA Model . . . . .	159
Topic 2: Incorporating Quality as a Separate Dimension in DEA . . . . .	164
Introduction . . . . .	164
Incorporating Quality into DEA Benchmarking . . . . .	165
Model I: Standard DEA Model . . . . .	167
Model II: Quality as an Output in a Standard DEA Model . . . . .	168
Model III: Independent Quality and Productivity Dimensions . . . . .	171
Model IV: Quality-Adjusted DEA: Q-DEA . . . . .	173
Q-DEA Benchmarking with Application to a Bank Branch Network . . . . .	176
Phase 1: Improve Branch Network Quality . . . . .	178
Phase 2: Use Q-DEA to Reduce Branch Network Operating Costs . . . . .	178
Results of Q-DEA Benchmarking . . . . .	181
Conclusion and Future Directions . . . . .	183
References . . . . .	184
<b>Part III Non-banking Financial Services</b>	
<b>12 Securities Market Applications: Risk Measurement of IPOs . . . . .</b>	<b>187</b>
Introduction . . . . .	187
Objectives . . . . .	189
Phase I: Comparable Selection . . . . .	190
Pool of Candidates . . . . .	192
Variable Selection . . . . .	193
Algorithm of Phase I . . . . .	198
Phase II: Short-Term Risk Assessment . . . . .	199
Stock Pricing Model . . . . .	200
Distribution of Stock Price 90 Days After the Issuing Day . . . . .	201

Calibrating the Distance Equation . . . . .	203
Validation of the Proposed Methodology . . . . .	204
Conclusion . . . . .	204
References . . . . .	205
<b>13 Securities Market Applications: Pension, Mutual and Hedge Fund Insights with DEA . . . . .</b>	<b>207</b>
Topic 1: Pension and Mutual Funds . . . . .	207
Introduction . . . . .	207
Background on Pension and Mutual Funds . . . . .	208
Pension Funds (PFs) . . . . .	208
Mutual Funds (MFs) . . . . .	210
Comparing Pension Funds and Mutual Funds . . . . .	210
Data and Variables . . . . .	211
Methodology: Directly Comparing PFs and MFs . . . . .	212
Results and Discussion: Directly Comparing PFs and MFs . . . . .	214
Considering All DMUs . . . . .	215
Combining Efficient DMUs . . . . .	217
Methodology: Bridging Pension Funds and Mutual Funds Indirectly . . . . .	218
Results and Discussion: Bridging Pension Funds and Mutual Funds Indirectly . . . . .	221
Conclusions . . . . .	223
Topic 2: Hedge Funds . . . . .	224
Introduction . . . . .	224
Examining Funds of Funds Type Hedge Funds . . . . .	225
Hedge Fund Performance . . . . .	225
Hedge Fund Efficiency . . . . .	226
Hedge Fund Strategies . . . . .	227
Concluding Remarks Regarding Hedge Funds and DEA . . . . .	229
References . . . . .	229
<b>14 Securities Market Applications: Stock Market Valuation of Securities and Financial Services – Insights with DEA . . . . .</b>	<b>233</b>
Introduction . . . . .	233
Topic 1: Stock Market Pricing Efficiency . . . . .	233
Topic 2: Private Firm Valuation . . . . .	240
Topic 3: Market Value Relationship to Corporate (Banking) Activity . . . . .	246
Introduction . . . . .	246
Topic 4: Stock Selection for Portfolios . . . . .	252
Introduction . . . . .	252
Conclusions . . . . .	256
References . . . . .	256

<b>15 Financial Services Beyond Banking: Credit Unions . . . . .</b>	<b>259</b>
References . . . . .	264
<b>16 Financial Services beyond Banking: Insurance . . . . .</b>	<b>265</b>
Introduction . . . . .	265
Topic 1: The Canadian Insurance Industry . . . . .	265
Insurance Models and Input/output Specifications . . . . .	266
Model I: Production Performance Approach . . . . .	267
Model II: Investment Performance Approach . . . . .	268
Analysis and Results . . . . .	269
Analysis by Insurer Characteristics . . . . .	274
Summary and Conclusions . . . . .	275
Topic 2: The Chinese Insurance Industry . . . . .	276
Conclusions . . . . .	280
References . . . . .	281
<b>17 Financial Services Beyond Banking: Corporate Failure</b>	
<b>Prediction . . . . .</b>	<b>283</b>
Conclusions . . . . .	308
References . . . . .	310
<b>18 Financial Services Beyond Banking: Risk Tolerance</b>	
<b>Measures for Portfolio Investors . . . . .</b>	<b>313</b>
References . . . . .	325
<b>Part IV Guidance on Applying DEA, Interpreting Results,</b>	
<b>Recognizing Caveats and Other Useful Information</b>	
<b>19 Guide to DEA Model Formulation . . . . .</b>	<b>329</b>
Introduction . . . . .	329
DEA Model Formulation: A Guide to Applying DEA	
to Evaluate and Manage Performance . . . . .	329
Objectives of the Analysis . . . . .	330
Operations of the Set of DMUs . . . . .	331
Defining Inputs and Outputs: Adequacy and Completeness	
of the Data . . . . .	333
Preliminary DEA Analysis: Testing the Reasonableness	
of the Results . . . . .	335
Using the Efficiency Scores: Limitations of Ranking . . . . .	337
Using the Information on Excess Resources	
and Excess Capacity . . . . .	338
Increasing the Power of the Analysis: Adjusting	
Constraints and Weights on Inputs and Outputs . . . . .	339
Impact of Other DMU Characteristics: Categorical Variables,	
Segmenting the Analysis, Quality . . . . .	340
Developing Best Practice Benchmarks . . . . .	341

Management of the Process: Converting DEA Results into Initiatives to Improve Performance . . . . .	341
Pitfalls and Roadblocks . . . . .	342
Results Interpretation (Graphs, Reports, Etc.) . . . . .	345
Conclusions . . . . .	350
References . . . . .	351
 <b>Part V Conclusions</b>	
<b>20 Conclusions and Recommendations . . . . .</b>	<b>355</b>
Reference . . . . .	356
 <b>List of DEA Software . . . . .</b>	<b>357</b>
 <b>About the Authors . . . . .</b>	<b>359</b>
 <b>DEA Books . . . . .</b>	<b>363</b>
 <b>Index . . . . .</b>	<b>367</b>

# List of Figures

Fig. 1.1	Radial improvement target (A') from CCR model for a 2-input and 1-output case .....	6
Fig. 1.2	Graphic representation of the five bank branches .....	31
Fig. 5.1	Profitability score distribution .....	82
Fig. 5.2	Productivity score distribution .....	84
Fig. 5.3	Comparison of efficiency scores for Country Red .....	85
Fig. 6.1	Sensitivity of spread ratio (scores from output wt. restricted/ unrestricted VRS models using all outputs) to permitted variation in AR constraints .....	94
Fig. 6.2	Number of efficient DMUs vs. permitted variation in AR constraints .....	94
Fig. 6.3	Convex envelopment surface defining DEA production possibility space – DMUs on blue hyperplanes are fully-efficient, those on red hyperplanes are weakly-efficient .....	97
Fig. 7.1	All branch types (A, B and C) use the same set of resources to provide all branch services used for the DEA analysis of Growth Bank's branch productivity. Each branch is using a different amount of each of the resources and offers all of the services. Each branch provides a different volume and mix of these services, depending upon its customer demand. Examples of branch types include urban, suburban, and shopping mall branches .....	104
Fig. 8.1	The theoretical, practical, and empirical frontiers .....	117
Fig. 8.2	Methodology to establish practical DEA frontier .....	118
Fig. 8.3	Comparison of DEA and P-DEA efficiency score distributions .....	121
Fig. 8.4	Input and output variables used in Tochaie 2003 .....	123
Fig. 8.5	CRS efficiency score distribution for all branches .....	123



Fig. 8.6	CRS efficiency score distribution for large branches .....	124
Fig. 8.7	CRS efficiency score distribution for small branches .....	124
Fig. 8.8	Distribution of the bank's WFI score for large branches .....	126
Fig. 8.9	DEA efficiency score distribution vs. the bank's WFI scores for large branches .....	127
Fig. 9.1	Potential input reduction at the current output level for Branch B .....	131
Fig. 9.2	Potential output enhancement at the current input level for Branch B .....	131
Fig. 9.3	Individual report for branch B .....	132
Fig. 9.4	Distribution of the scores obtained from the second stage, overall model .....	137
Fig. 10.1	Branch operational efficiency model from Paradi et al. (2010) .....	150
Fig. 10.2	Branch profitability model from Paradi et al. (2010) .....	151
Fig. 10.3	Churn model efficiency distribution .....	156
Fig. 10.4	Aggregate market model efficiency distribution .....	157
Fig. 11.1	Distribution of client service ratio by branch size group .....	161
Fig. 11.2	Distribution of throughput ratio by branch size group .....	162
Fig. 11.3	Comparison of DEA efficiency and client service ratio .....	162
Fig. 11.4	Comparison of DEA efficiency and the bank's existing customer satisfaction benchmark .....	163
Fig. 11.5	Efficient frontier (all branches service 1,000 transactions), where A(100) = Branch A with quality rating = 100 .....	167
Fig. 11.6	Quality-productivity branch distribution – high and low quality and productivity quadrants .....	172
Fig. 12.1	General layout of the DEA model .....	193
Fig. 12.2	Several potential improvement directions for DMU E .....	197
Fig. 13.1	Snapshot of theoretical methodology .....	220
Fig. 14.1	Quarterly Treynor measure for software portfolios .....	239
Fig. 14.2	DEA inputs and outputs of modified valuation model from Anadol et al. (2014) .....	241
Fig. 14.3	Bank intermediation efficiencies, single DEA analysis on the entire data sample .....	249
Fig. 14.4	Bank intermediation efficiencies, DEA analysis using 5-year windows .....	250
Fig. 14.5	Bank production efficiencies, DEA analysis using 5-year windows .....	251
Fig. 15.1	Difference of mean scores between healthy and failed credit unions over time from different models, assets greater than \$2 million .....	262

Fig. 16.1	Model I – inputs and outputs included in the production performance model .....	268
Fig. 16.2	Model II – inputs and outputs included in the investment performance model .....	269
Fig. 16.3	Unadjusted and adjusted mean risk management efficiency of Chinese insurers .....	279
Fig. 16.4	Unadjusted and risk-adjusted mean efficiencies of Chinese insurers .....	279
Fig. 16.5	Differences in mean efficiencies between SOEs and non-SOEs .....	280
Fig. 17.1	Current limitations of DEA and other methodologies in bankruptcy prediction .....	286
Fig. 17.2	Bankrupt and non-bankrupt classification accuracy 1-year prior .....	289
Fig. 17.3	Total classification accuracy comparison between Altman and DEA (SBM) .....	291
Fig. 17.4	Illustration comparing regular ( <i>left</i> ) and Negative ( <i>right</i> ) DEA .....	292
Fig. 17.5	Type I error from Z-score by years prior to bankruptcy .....	301
Fig. 17.6	Variation of classification and error rates by cut-off layer from IS model, up to 1 year prior to bankruptcy .....	303
Fig. 17.7	Comparison between layering and non-layering techniques – 1 year prior to bankruptcy .....	306
Fig. 17.8	Distribution of second-stage layered scores .....	307
Fig. 17.9	Probability of bankruptcy as a function of layered score .....	309
Fig. 18.1	Comparison of DEA and FinaMetrica scores for all clients only .....	318
Fig. 18.2	Comparison of DEA and FinaMetrica scores for all subjects .....	319
Fig. 18.3	Quadratic fit of average risk tolerance vs. age .....	321
Fig. 19.1	Commercial bank branch DEA model .....	346
Fig. 19.2	Individual results: Branch 78 score = 0.91 .....	346
Fig. 19.3	Comparison chart to benchmark: Branch 6 cost-efficiency = 0.78 .....	347
Fig. 19.4	Efficiency and asset size in two models of the Canadian life and health insurance industry .....	348
Fig. 19.5	Insurer ownership type and efficiency .....	349
Fig. 19.6	Comparison between client results and DEA .....	349
Fig. 19.7	Portfolio types and their efficiency in earnings .....	350

# List of Tables

Table 1.1	Timeline of DEA banking applications .....	27
Table 1.2	Illustrative example of five bank branches .....	28
Table 1.3	DEA results for five bank branches .....	32
Table 1.4	Inefficiency in branch B2 calculated by DEA .....	33
Table 1.5	Multiplier form of DEA mathematical model .....	35
Table 4.1	Descriptive statistics for inputs and output in 2006 and 2007 (in billion rupiahs) .....	72
Table 4.2	Descriptive statistics of the DEA efficiency measures, 2006 and 2007 .....	73
Table 4.3	Banking data of commercial banks in India as of June 1998 .....	74
Table 4.4	Descriptive statistics of efficiency scores by bank ownership .....	74
Table 4.5	Key questions regarding stakeholder views from Avkiran and Morita (2010) .....	76
Table 5.1	Profitability model data – means, in USD .....	80
Table 5.2	Productivity model data – means .....	81
Table 5.3	Intra-country profitability model results .....	82
Table 5.4	Inter-country profitability model results .....	82
Table 5.5	Intra-country productivity model results .....	83
Table 5.6	Inter-country productivity model results .....	84
Table 6.1	Inputs of production model .....	89
Table 6.2	Average efficiency scores of the branch system .....	90
Table 6.3	Data statistics, standard times and average salaries .....	91
Table 6.4	Results for DEA models .....	93
Table 6.5	Comparison of overall and within group DEA results: all outputs, VRS .....	95

Table 6.6	Summary of normalized data for small urban branches .....	98
Table 6.7	Efficiency results of technically inefficient branches .....	98
Table 6.8	Efficiency results of technically but not scale efficient branches .....	98
Table 6.9	Summary of input-oriented efficiency results for small urban branches .....	99
Table 7.1	Growth Bank branch productivity ratings .....	106
Table 7.2	Growth Bank, potential resource savings in less productive branches .....	107
Table 7.3	Potential service volume expansion .....	108
Table 8.1	Example of a DEA benchmark for an inefficient unit, i.e. DMU#12 .....	115
Table 8.2	Optimal DEA input weights for DMU #12 .....	116
Table 8.3	New DEA benchmark determined for an inefficient DMU#12 by prioritizing <i>personnel</i> reduction, i.e. DMU#12' .....	116
Table 8.4	Data statistics from Sowlati and Paradi 2004 .....	119
Table 8.5	Input and output comparisons for original and newly generated DMU #23 .....	120
Table 8.6	Comparison of inputs, outputs and P-DEA efficiency scores for real and artificial units .....	120
Table 8.7	Results of changing input and output bounds and $\delta$ .....	122
Table 8.8	Summary of CRS and VRS DEA mean efficiency results by geographical area .....	125
Table 9.1	Individual report for branch B .....	133
Table 9.2	Comparison of regular and handicapped DEA results, overall and by bank .....	136
Table 9.3	Reference vectors for input/output vectors .....	140
Table 9.4	Statistical descriptions of groups based only on group leaders .....	141
Table 9.5	Comparison of within group referencing of inefficient DMUs .....	142
Table 10.1	Annual and average corporate index scores for largest Canadian banks and trust companies .....	147
Table 10.2	Simulation results summary with $RI = 1.5$ .....	149
Table 10.3	Spearman's rank correlation between the true and estimated efficiencies .....	150
Table 10.4	Summary of branch efficiencies from basic, CA and NC-DEA models .....	151
Table 10.5	"Component" market model .....	153
Table 10.6	"Aggregate" market model .....	154
Table 10.7	Churn model .....	154

Table 10.8	Delta model .....	155
Table 10.9	Cluster statistics .....	155
Table 11.1	CRS vs. VRS results DEA results for all branches .....	161
Table 11.2	DEA customer satisfaction results for branch-hour DMUs .....	163
Table 11.3	Bank branch example .....	166
Table 11.4	Model I – benchmarking productivity with DEA <i>excluding quality</i> .....	169
Table 11.5	Model II – benchmarking with quality as an output .....	170
Table 11.6	DEA productivity ratings .....	174
Table 11.7	Q-DEA benchmarking .....	175
Table 11.8	Branch data used for Q-DEA benchmarking .....	177
Table 11.9	Q-DEA benchmarking applied to a US branch network .....	179
Table 11.10	Q-DEA benchmarking distribution of productivity ratings in Phase 2 in the US bank application .....	180
Table 11.11	Potential savings identified with Q-DEA and actual resource savings realized within 6 months of completing the Q-DEA study .....	181
Table 13.1	Inputs and outputs for DB plans and Combo plans .....	212
Table 13.2	Inputs and outputs for DC plans .....	212
Table 13.3	Considering all DB, Combo and MFs for VRS, ND-VRS and MV-DEA models .....	215
Table 13.4	Considering all DC and MFs for VRS, ND-VRS and MV-DEA models .....	216
Table 13.5	Combining efficient DB, Combo and MF DMUs for VRS, ND-VRS and MV-DEA models .....	217
Table 13.6	Combining efficient DC and MF DMUs for VRS, ND-VRS and MV-DEA models .....	218
Table 13.7	Theoretical classification of pension plans .....	220
Table 13.8	Results for DB, Combo and DC plans .....	223
Table 13.9	Input and output variables for the VRS hedge fund model .....	225
Table 13.10	Input and output variables for hedge fund model .....	227
Table 13.11	List of hedge fund strategies .....	228
Table 13.12	Potential input and output variables .....	228
Table 14.1	DEA pricing efficiency model variables from Tam (2001) .....	238
Table 14.2	Summary of inverse of DEA efficiency scores from Tam (2001) .....	238
Table 14.3	Market cap. estimate and upper bound for Cheniere Energy .....	243

Table 14.4	Distance indicators and MC ranges for Cheniere Energy and its peers .....	244
Table 14.5	Lower bound MC determination for Costco .....	245
Table 14.6	Model variables in production model .....	247
Table 14.7	Model variables in intermediation model .....	248
Table 14.8	Change in results from adding total or excess return as an additional output to DEA window analysis models .....	252
Table 14.9	Quarterly returns for the 22 portfolios .....	255
Table 15.1	Mean failure prediction index and standard deviation for years prior to failure .....	263
Table 16.1	Number of insurers based on their characteristics .....	270
Table 16.2	Average efficiency scores and statistical tests of efficiency differences .....	271
Table 16.3	DEA results – production performance model .....	272
Table 16.4	DEA results – investment performance model .....	273
Table 16.5	Efficiency comparison and statistical tests on subsets of insurers .....	275
Table 16.6	Variables used by Huang and Paradi (2011), along with descriptive statistics .....	278
Table 17.1	Confusion matrix for prediction outcomes .....	284
Table 17.2	Non-negative input and output variables .....	287
Table 17.3	Number of companies in Groups 1 and 2 .....	288
Table 17.4	Cut-off points for SBM model .....	289
Table 17.5	Classification accuracies of Group 2 firms .....	290
Table 17.6	Summary of DMUs in data samples .....	293
Table 17.7	Corporate performance indicators identified in the literature .....	293
Table 17.8	Variables identified as potential inputs and outputs .....	294
Table 17.9	Average efficiency scores for bankrupt and non-bankrupt firms in normal DEA models, with optimal cut-off values and the corresponding classification accuracies .....	294
Table 17.10	Average efficiency scores for bankrupt and non-bankrupt firms in (output-oriented) Negative DEA models, with optimal cut-off values and the corresponding classification accuracies .....	295
Table 17.11	Classification accuracies for Negative DEA model #3 using the layering technique .....	296
Table 17.12	Out of sample (i.e. 1996 data) classification accuracies from combining NDEA3 and DEA5 models, using layering .....	296

Table 17.13	Input and output variables of IS, BSA and BSL (financial) DEA models .....	298
Table 17.14	Average median ratio values by firm state .....	299
Table 17.15	Managerial decision-making (MDM) variables .....	300
Table 17.16	Market and economic (ME) factor models .....	300
Table 17.17	Summary of first-stage results for IS, BSA, BSL and MDM models .....	302
Table 17.18	Correlations between first-stage DEA scores .....	302
Table 17.19	Cut-off layer, and type I error, type II error and accuracy rates for first-stage models .....	303
Table 17.20	Second-stage model predictions with classifications by zones .....	304
Table 17.21	Correlation of first-stage models' layered scores .....	305
Table 17.22	Error from classification by layering of second-stage model and individual first-stage models .....	305
Table 17.23	Performance comparison of layering and non-layering techniques .....	306
Table 17.24	Probabilities of bankruptcy (B) and non-bankruptcy (NB) by layer number .....	308
Table 17.25	Classification by layering and fitted second order 1 year prior to bankruptcy probability polynomials for different windows .....	309
Table 18.1	Demography of risk tolerance .....	315
Table 18.2	Data statistics for all respondents and sample of clients only .....	317
Table 18.3	Summary of results from the SBM DEA model .....	317
Table 18.4	Variation of average risk tolerance with education level .....	320
Table 18.5	Variation of average risk tolerance with income level .....	320
Table 18.6	Variables used in first-stage models in Cooper et al. (2014) .....	323
Table 18.7	Results from first- and second-stage models .....	323
Table 18.8	Comparison of risk tolerance scores by gender .....	324
Table 19.1	DEA overcomes these issues that other methods lack .....	343
Table 19.2	Regional comparisons .....	347

# Introduction

Data envelopment analysis was first titled with this name in the paper by Charnes, Cooper, and Rhodes in 1978. The initials DEA have since been widely adopted. The concept was previously exposed in Farrell's seminal paper (1957): "The measurement of productive efficiency." Farrell did not have the power of modern computing equipment at his disposal, so the development of practical applications was not feasible in a practical sense. But time passed and technology developed so that Farrell's work became possible to apply to complex problems with multiple inputs and outputs. Linear programming capabilities allowed the DEA models to be used for varied problems. Running DEA often required rerunning a linear program thousands of times, a capability that was not readily available in the 1950s. Today, running numerous linear programming iterations required for DEA can be done on the average personal computer by simply using DEA custom-coded programs or even Microsoft Excel.

Slowly, researchers in operational research and economics began to apply DEA to their problems. With few exceptions, their primary goal was to extend the theoretical foundations of the science and report this in traditional academic refereed journals in management science, economics, social science, and mathematics. As more researchers became involved in looking at DEA as a fruitful approach to management and economic problems and their works were published, the literature grew, at first slowly and in recent years quite rapidly. While in the early days it was possible to keep up with the new papers as they appeared (e.g., Seiford 1997; Emrouznejad et al. 2008), this is now essentially impossible as it would take a person working full time just to assemble the bibliography. The number of books alone now published is around 100 and growing. The DEA technology is now well established but still developing, and relatively small theoretical additions, extensions, and refinements continue to be reported in the academic literature. One of the best sources of the most up-to-date information on DEA is found at A. Emrouznejad's DEA Zone on the web (2017): [www.deazone.com](http://www.deazone.com).



However, the major challenge and unfinished DEA work is, in our view, that only a small portion of the published works deal with applications of DEA to real-life problems and even fewer result in production systems making use of DEA. The reader might challenge this assertion, so let us clarify. While most recent papers used real data obtained from credible sources, such as the OECD, national statistics from agriculture to retirement homes, and the financial and economic data sources, studies based on such data do not enable managers to obtain directions on how to enact policy or improve practice in their businesses or other organizations. Even when there is the potential to apply the DEA findings to real operating organization datasets, the results of the analyses are published without pursuing the application to generate the potential benefits. While there are examples where the results have been applied and the positive and negative results are reported, these papers reflect an incredibly small fraction of the total DEA published literature. There are applications that have been successful that have not been published, and while we cannot know the universe of the works not published, discussions with academics and end users of DEA suggest that these unpublished applications are not likely to be very large in number. Two fields that stand out in these studies are health care and banking where hundreds of papers were written over the past couple of decades, but with very few being of practical use to the people who operate these institutions.

The early focus of DEA was applying it to units in any organization that have control over their activities, and where there is some manager that assesses performance of each unit and makes decisions about how the unit operates in an effort to improve its outcomes. The term adopted for these operating units was *decision-making units*, or DMUs. These initials, DMUs, are well understood by the DEA community, but this is not a term that is naturally found or used in business, government, or other organizations. The terminology in itself may be sufficiently arcane and unfamiliar to potential users that it may have contributed to the slow adoption of DEA. The current use of DEA continues to heavily focus on understanding and improving the performance of the defined DMUs, but has also broadened to recognize DEA's ability to identify relationships in complex operating data that offer new insights into the way organizations operate and other paths to manage performance (Sherman and Zhu 2013).

The definitions of what a DMU is determine the usability of the results. For economists, the aggregate data is useful when they advise governments on policy or evaluate the national or international health of certain sectors of interest. The DMU may be defined as a political unit, country, industry, etc. But useful direction for the managers of units such as bank branches, hospital departments, farms, retirement homes, etc. is seldom provided, yet this is where real operating benefits can be achieved. For example, when a study is conducted on the efficiency, productivity, or effectiveness of the banking industry, the outcomes for each bank (the DMUs in the models) offer no implementable findings as the data is aggregated and applies to the DMU as a whole. Of course, the outcome of such a study may well be useful for the regulator or government evaluation of the health of the industry and in identifying regulatory policies that would improve overall productivity. A concrete example of DEA being applied to help regulate an industry is the utilities sector,

where it has been used to manage electricity producers in Europe and Brazil (Agrelle et al. 2005).

When the focus is on DMUs that are finite operating units such as bank branches, clinics, physicians, hospitals, nursing homes, and focused services in health care, DEA provides, in addition to an assessment of the DMUs, insights that can allow a manager to directly adjust methods of operations. These adjustments can provide the opportunity to measurably improve the performance of the DMUs analyzed with DEA.

This book is intended to address the challenge of how to apply the DEA technology to data, where the data is relevant and detailed enough to allow results to be useful to the managers by implementing the outcomes from the study to improve the performance of their organizations. In other words, we look at the practitioners' problem of applying improvements to the businesses or institutions where the benefits are directly received by the owners, employees, and/or customers of the firm. Of course, the entire firm benefits from the individual improvements. For example, an analysis of a retail chain store or franchising operation where managers do have the power to implement the improvements suggested by a DEA analysis could result in lower costs for individual operating units (the DMUs in this type of study), improving profitability of these units and thus an augmentation of the system-wide success and attractiveness of owning one of the franchised units.

Our intent and objective is to provide any reader of this book a set of useful approaches and techniques which they can apply and, if done as suggested in this volume, would enable the reader to improve their firm's performance (or that of their client firm if they are consulting for them). However, there are many sectors in a large economy and no single book can cover them all. Therefore, we restricted ourselves to the financial sector where there are a number of studies published examining the actual performance level of the firm and where the firm should go to reap the benefits of the study. Perhaps it would be appropriate to see this book as a *how-to* manual where the practitioner or analyst can find a study that relates to their problem, often directly, while other times they may find an example where there are similarities to their organization but which requires some adaptation to be effective.

## References: Introduction

- Agrelle, P.J., Bogetoft, P., Jørgen, T: DEA and dynamic yardstick competition in Scandinavian electricity distribution. *J. Product. Anal.* **23**(2), 173–201 (2005)
- Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2**(6), 429–444 (1978)
- Emrouznejad, A.: Ali Emrouznejad's data envelopment analysis. <http://www.deazone.com> (2017). Accessed 30 Nov 2016
- Emrouznejad, A., Parker, B.R., Tavares, G.: Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Econ. Plan. Sci.* **42**(3), 151–157 (2008)

- Farrell, M.J.: The measurement of productive efficiency. *J. R. Stat. Soc.* **120**(3), 253–290 (1957)
- Seiford, L.M.: A bibliography for data envelopment analysis (1978–1996). *Ann. Oper. Res.* **73**, 393–438 (1997)
- Sherman, H. D., Zhu, J.: Analyzing performance of service organizations. *MIT Sloan Manag. Rev.* **54**(4), 37–44 (2013)

# **Part I**

## **Data Envelopment Analysis, in Brief, with Little Math!**

### **Introduction**

In Part I of this book we provide an overview of DEA but without a lot of mathematics, except to allow the explanations to make sense. Our intent is to allow the reader to assess the technology and understand it well enough to delve into whatever details he or she feels necessary to their needs. This part also provides a brief survey and summary of some of the large body of published DEA studies on banking and other financial services.

# Chapter 1

## DEA Models Overview

We begin with the basic DEA Models and some useful extensions (although we expect that some will see it as too much while others as too little). While we promised to minimize the mathematics, some are, unfortunately, unavoidable. We have excluded any specific discussion of the underlying linear programming (LP) mathematics that drives DEA, and while some general understanding of this is helpful for understanding the academic literature, it is not needed to understand the benefits and ways to apply DEA.

The next issue here is how to select what are “inputs” and what are “outputs”. One would assume that this is easy since whatever is used in the production model is an input and what is produced is the output. But there are some issues, such as undesirable outputs (e.g. bad loans) and inputs where we might want more (e.g. deposits in a bank branch). To make matters more confusing, some measures may well be used as inputs in some models while outputs in others – such as bank deposits. We also address some data issues in this part.

Model formation is another subject fraught with controversy and we point to some issues and suggest ways to address these problems.

Finally, we provide a brief history of DEA and its development and sketch out the DEA models family tree to show how things connect together from the Charnes et al. (1978) model to some of the more sophisticated models that have been developed since that seminal paper saw the light of day. Some of the milestone applications of DEA in the financial services are also provided.

## Basic DEA Models

### *Introduction*

Data Envelopment Analysis (DEA) is an example of a frontier approach. Frontier approaches identify and assess the areas or examples of best performance or best practice within the sample, i.e. those located on the “frontier”. The frontier identified by DEA suggests the best performance within the group of operating units being evaluated and does not promise or even suggest that these represent the theoretically best performance. These methods can be contrasted with regression techniques that seek to explain the average behavior within a sample. Frontier techniques can be divided into two types: parametric and non-parametric. Parametric techniques specify a frontier function to be fitted to the data, with or without accounting for noise in the data. DEA is a non-parametric approach, meaning that no prior functional form is assumed for the frontier, outside of a simple assumption of piecewise linear connections of units on the frontier. The ability to apply DEA without assuming a functional form is a very powerful characteristic. This means that the analysis can proceed without knowing the production function, which is the way inputs are transformed into outputs. Non-parametric approaches can simultaneously handle multiple inputs and outputs, but do not account for noise in the data, treating all deviations from the frontier as inefficiencies (Cummins and Zi 1998).

As an efficiency measurement and evaluation methodology, DEA is particularly useful in cases where sample units, termed decision making units (DMU), use multiple inputs and outputs, and are operating under comparable conditions. DEA primarily measures technical efficiency, i.e. focusing on levels of inputs relative to outputs, as opposed to economic efficiency which would also consider market prices. The use of levels of inputs and outputs is another powerful characteristic of DEA, in that it can incorporate inputs and outputs in the natural units in which they are measured and does not require them to be converted to the same units of measure – specifically, they are not required to be converted to monetary units.

DEA permits the evaluated DMUs to appear to be as good as possible, a feature that can be deemed as providing a “fair” evaluation of the DMUs in the sense that the analysis should limit objections amongst DMUs regarding their evaluations. This characteristic stems from the optimization underpinning of DEA, where DEA assigns the highest efficiency rating to each DMU compared with the set of DMUs being analyzed. It essentially gives the “benefit of the doubt” to each unit. From a management perspective, DEA will be less likely to erroneously identify an efficient unit as inefficient, and while it may not capture all inefficient units, the ones identified as inefficient will have real potential for improvement.

On the contrary, this same “fairness” can permit DMUs to select evaluation criteria that may be deemed as inappropriate or unrealistic. The DEA results, including the evaluation criteria, can be reviewed and adjusted by the user to rerun DEA to include more appropriate criteria. This is one of several ways DEA can be adapted to the specific operating environment of the DMUs. These

adaptations make DEA more powerful but also require that the user understand the nature of these added constraints and how that nature affects the way the DEA results are analyzed.

Another advantage of DEA is that it suggests explicit improvement targets for inefficient DMUs, namely the *benchmark* or point on the frontier to which it is being compared in order to measure its efficiency. Furthermore, this frontier point will be defined as the linear combination of one or more actual DMUs that are efficient (i.e. on the efficient frontier). The inefficient DMU is presented with a relevant set of efficient DMUs, called its *reference set* (sometimes referred to as the efficient reference set). The reference set represents the specific efficient DMUs against which the inefficient DMU is judged to be inefficient, and changes to improve the inefficient DMU can be most directly determined by analyzing differences between the inefficient DMU and its reference set. The unit managers thus receive actionable advice that is perceived by them as fair and equitable. Identifying the amount of excess resources consumed or potential increase in outputs possible in inefficient units compared to the DMUs in the efficient reference set may be the most powerful and useful feature of DEA. This perspective offered by DEA is unique, in that it is not provided by any other method known to the authors.

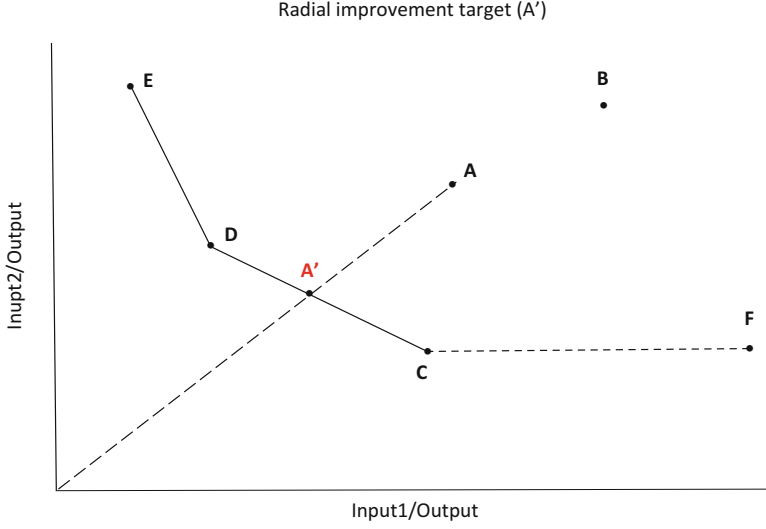
If one were to reread these introductory paragraphs, the clear implication is that DEA is an extremely powerful analytic and management tool. We believe it has been underutilized and hope this volume will open the path to greater utilization. At the same time, we emphasize that DEA is a complement to operating and financial analytical tools, and is not offered as a replacement or a method that must be used exclusively for enhancing business operations.

## ***Model Types***

There are three types of basic DEA models: radial, additive and slack-based measure models. These models are detailed in the following sections, along with a discussion of their properties.

### **Radial Models: CCR**

The original DEA model proposed by Charnes et al. (1978), also termed the CCR model, was a radial model. In such a model, a DMU's efficiency score is derived from the extent to which all of its inputs can be contracted and/or its outputs expanded, where this contraction or expansion occurs proportionately. For example, in the case of a model seeking to reduce inputs, the greatest percentage reduction in all inputs is sought; hence the term "radial", as the examined input possibilities occur on the line extending radially from the origin of the input space



**Fig. 1.1** Radial improvement target ( $A'$ ) from CCR model for a 2-input and 1-output case

(i.e. zero values for all inputs) to the DMU's original inputs (c.f. DMU A in Fig. 1.1).

The CCR model assumed that the production technology, also called the production possibility set, exhibited constant returns to scale (CRS). Charnes, Cooper and Rhodes gave the model in two orientations: input- and output-orientations. The orientation corresponds to the viewpoint taken in improving the inefficient units, whether the goal would be to reduce excess inputs consumed or expand shortfalls in outputs produced, respectively, to move the inefficient unit to the frontier. The frontier – sometimes referred to as the best practice frontier – in Fig. 1.1 is composed of the lines joining points EDC. The line joining CF is termed a *weakly efficient skirt* of the frontier, since points on this line – other than C – will have a radial efficiency score of one, but could still reduce Input 1 without adversely affecting other variables, see Eq. 1.3 below. For a model with  $m$  inputs variables,  $s$  output variables, and  $n$  DMUs, the envelopment form of the input-oriented model is given by Cooper et al. (2007):

$$\begin{aligned}
 & \min_{\theta, \lambda} \theta \\
 & \text{subject to} \quad \theta x_o - X\lambda \geq 0 \\
 & \quad Y\lambda \geq y_o \\
 & \quad \lambda \geq 0,
 \end{aligned} \tag{1.1}$$

where  $x_o$  and  $y_o$  are the column vectors of inputs and outputs respectively for DMU<sub>o</sub>,  $X$  and  $Y$  are the matrices of input and output vectors respectively for all DMUs,  $\lambda$  is the column vector of intensity variables denoting linear combinations of DMUs, and the objective function  $\theta$  is a radial contraction factor that can be applied



to  $DMU_o$ 's inputs. As DEA measures efficiency *empirically* relative to the data sample, having too few DMUs will generally result in a large proportion of them being found to be efficient. A general rule of thumb as to the minimum number of DMUs in relation to the number of variables to have a meaningful result with a clear set of efficient and inefficient units is given by Banker et al. (1989):

$$n \geq \max\{m \times s, 3(m + s)\}, \quad (1.2)$$

where  $m$ ,  $s$  and  $n$  are the numbers of inputs, outputs and DMUs respectively. This is more a rule of thumb than a rule, which by its nature is a qualitative judgment. When using DEA, it is possible to get very useful results with fewer DMUs than are suggested by this guideline as long as the results are analyzed understanding that there is a small sample of DMUs and thus limited discriminatory power in the model.

The model given in Eq. 1.1 seeks to identify the largest proportion by which all inputs can be reduced (i.e.  $1-\theta$ ), while at least producing the same level of outputs as the original DMU. Also note that it is assumed that all linear combinations of two or more actual DMUs also represent possible productions, i.e. combinations of inputs and outputs. The linear programming (LP) optimization given in Eq. 1.1 is repeated for each DMU. The optimal value of  $\theta$ , denoted by  $\theta^*$ , obtained can be considered the efficiency score of the DMU in question, and this value will range from zero to one, inclusive. (Frequently, the 0–1 scores are reported as percentages – 0% to 100%.) Efficient units will not be able to further reduce inputs and hence have an efficiency score ( $\theta^*$ ) of one. The efficient unit with a rating of 1 or 100% is relatively efficient compared to the DMUs in the study and is not represented as having reached absolute efficiency in an engineering of theoretical sense.

In some instances, it may be possible to further improve the DMU's production performance after the radial optimization. For example, in an input-oriented model, it may be possible to reduce the usage of the first input to 80% of the initial amount, while only reducing the remainder of the inputs to 85%. The  $\theta^*$  would be 0.85, but in suggesting an improvement target for the DMU, it would be more intuitive to incorporate the additional possible improvement in the first input. Similarly, even though the input-oriented model focuses on reducing inputs, it may be possible to produce more outputs using the same amount of inputs. These additional possible input reductions and output expansions are termed *slacks*, and can be optimized through a second stage to the DEA model, which is given as Eq. 1.3 (Cooper et al. 2007):

$$\begin{aligned} \max_{\lambda, s^-, s^+} \quad & \omega = e_m s^- + e_s s^+ \\ \text{subject to} \quad & s^- = \theta^* x_o - X\lambda \\ & s^+ = Y\lambda - y_o \\ & \lambda \geq 0, s^- \geq 0, s^+ \geq 0, \end{aligned} \quad (1.3)$$

where  $e_m$  and  $e_s$  are row vectors of  $m$  and  $s$  ones respectively,  $s^-$  and  $s^+$  are column vectors of input and output slacks respectively,  $\theta^*$  is the optimal input contraction

obtained from the first stage (Eq. 1.1), and the remaining variables are as previously described.

Equation 1.1 is termed the *envelopment* form of DEA. The same model can be presented in another, equivalent model, termed the *multiplier* form (Cooper et al. 2007):

$$\begin{aligned} \max_{v,u} \quad & uy_o \\ \text{subject to} \quad & vx_o = 1 \\ & -vX + uY \leq 0 \\ & u \geq 0, v \geq 0. \end{aligned} \tag{1.4}$$

As with Eq. 1.1, the multiplier form of the DEA model is run once for each DMU in the sample. The model selects virtual or marginal weights for the input and output variables,  $v$ 's and  $u$ 's respectively, in such a way as to maximize the efficiency score of the DMU<sub>o</sub>, where efficiency is measured as the ratio of the virtual output (i.e. sum of outputs weighted by the virtual weights) to the virtual input. The only restriction on the chosen weights are that they be non-negative and feasible for the sample, i.e. that applying the same weights to any DMU in the sample will not produce an efficiency score greater than one. It is this interpretation of the multiplier form of DEA that lends to the prior assertion of the fairness of DEA models to the evaluated DMUs.

Note that the above model allows the weights,  $u$  and  $v$  to be greater than or equal to zero. The intention is that the weights should be greater than zero, as allowing a weight to be zero effectively eliminates that input or output from the assessment of a DMU. For computational and other reasons, some DEA programs allow zero weights. Some may use a very small minimum value to at least include all inputs and outputs in the assessment of every DMU in the dataset. If one uses a DEA program and there are zero weights, the interpretation of the results should explicitly consider the implications of the zero weights, as each DMU can look relatively more efficient by removing the inputs/outputs that it tends to use/produce least efficiently via assigning zero weight to those inputs and outputs. Most commercial DEA software will run both envelopment and multiplier forms of the models, as well as any second-stage slack optimizations.

The envelopment and multiplier forms of the model form a primal-dual pair of LPs, and as such the optimal solutions to Eqs. 1.1 and 1.4, and thus the determined efficiency scores, will be the same. Any LP problem, termed the *primal*, can be transformed, through a set procedures known as *taking the dual* (c.f. Appendix A, Cooper et al. 2007) into another LP, the *dual*, and the optimal solutions to each of the two will be the same, provided a solution exists. The primal-dual terminology is non-specific, since taking the dual of the dual program retrieves the original primal program. As such, each of two could be considered the primal or the dual. Hence, this book will avoid the labels *primal* and *dual* and instead employ the more descriptive and specific labels of *envelopment* and *multiplier* forms of DEA. It should be noted that most DEA studies tend to refer to the multiplier form as the primal LP model.

The envelopment form of the output-oriented CRS radial model is presented as Eq. 1.5 (Cooper et al. 2007).

$$\begin{aligned} \max_{\phi, \lambda} \quad & \phi \\ \text{subject to} \quad & x_o - X\lambda \geq 0 \\ & Y\lambda \geq \phi y_o \\ & \lambda \geq 0, \end{aligned} \tag{1.5}$$

where  $\phi$  is the radial expansion factor that can be applied to DMU<sub>o</sub>'s outputs.

Analogous to the input-oriented version of the model, the model seeks the maximum factor by which all outputs can be simultaneously expanded. Taking the inverse of the optimal expansion factor,  $\phi^*$ , produces an efficiency score in the standard sense, i.e. ranging from zero to one. Efficient units will not be able to increase outputs produced from the same inputs, and thus have a  $\phi^*$  and efficiency score of one. One property of radial CRS DEA models is that the efficiency scores determined for DMUs are the same in both input- and output-oriented models, i.e.  $\theta^* = 1/\phi^*$ . For the multiplier form of the output-oriented CCR model, refer to Cooper et al. (2007).

*Appendix – Basic DEA model illustration:* The DEA model assuming no knowledge of linear programming is explained and applied to a simple dataset in the appendix to this chapter. This fundamental description of DEA illustrates the way it might be used to identify best practice DMUs, inefficient DMUs, and the potential benefits if inefficient DMUs become as efficient as the efficient DMUs by making the changes suggested by DEA.

## Radial Models: BCC

Banker et al. (1984) developed a radial DEA model where the production technology exhibits variable returns to scale (VRS). The envelopment form of the input-oriented version of the model is given by Cooper et al. (2007):

$$\begin{aligned} \min_{\theta_B, \lambda} \quad & \theta_B \\ \text{subject to} \quad & \theta_B x_o - X\lambda \geq 0 \\ & Y\lambda \geq y_o \\ & e_n \lambda = 1 \\ & \lambda \geq 0. \end{aligned} \tag{1.6}$$

Comparing Eqs. 1.1 and 1.6, it can be seen that they differ in the addition of a constraint that the sum of the intensity variables,  $\lambda$ 's, be equal to one in the VRS model. The effect of this constraint is to limit a DMU to being compared to other

DMUs that are of roughly the same operational scale, which allows for the existence of VRS, i.e. increasing, constant or decreasing returns to scale (RTS). The CRS efficiency score will be less than or equal to the VRS score, and the ratio of CRS/VRS scores gives a measure of the DMU's scale efficiency, i.e. the effect on its productivity from potentially not operating at the optimal scale. This relationship between CRS and VRS scores holds for all DEA models.

Varying the constraint on the sum of  $\lambda$ 's to being less than or equal to one results in a non-increasing returns to scale (i.e. permitting constant or decreasing RTS) model. Restricting the sum of  $\lambda$ 's to be greater than or equal to one yields a non-decreasing RTS model. The effect of these constraints on the  $\lambda$ 's affects the RTS properties of other DEA models in the same manner.

The VRS model is frequently applied and can offer useful additional insights to those obtained from a CRS model on the same dataset. When the VRS or CRS model is specified, the reason for choosing one over the other should also be noted. One of the advantages of DEA is that one need not know the functional form, which would include knowing the returns to scale characteristics. There are also cases where a larger unit is less efficient than a smaller unit, and analyzing this situation where there are expected to be increasing returns to scale can overlook the real possibility that the large unit is less efficient due to the way it operates and not due to decreasing returns or any scale effect. Applying both CRS and VRS would help identify the inefficiency in the larger unit.

The reader is referred to Cooper et al. (2007) for the multiplier form of the input-oriented BCC model, as well as the two formulations for its output-oriented version. Note that unlike the case for CRS models, it is not generally the case that the efficiency scores from the input- and output-oriented versions of VRS models will be the same.

In some analysis situations, there may not be an intuitive reason to emphasize either input reduction or output maximization, and instead it may be reasonable to pursue both. To address this situation, radial DEA models can express in a non-oriented form. The CRS version of the envelopment form of the non-oriented radial model is (Tam 2004):

$$\begin{aligned}
 & \max_{g, \bar{\lambda}} \quad g \\
 & \text{subject to} \quad x_o - X\bar{\lambda} \geq 0 \\
 & \quad Y\bar{\lambda} \geq gy_o \\
 & \quad g \geq e_n \bar{\lambda} \\
 & \quad \bar{\lambda} \geq 0,
 \end{aligned} \tag{1.7}$$

where  $g = \varphi/\theta$ ,  $\bar{\lambda} = \lambda/\theta$ , and the other variables are as defined in Eqs. 1.1, 1.5, and 1.6. In this model both a radial contraction,  $\theta$ , and radial expansion,  $\varphi$ , are determined, subject to constraints that the target to which a DMU is being compared cannot use more inputs ( $\theta \leq 1$ ) or produce less outputs ( $\varphi \geq 1$ ). The efficiency score for the DMU is given by  $1/g^* = \theta^*/\varphi^*$ . For the CRS model, the efficiency score from the non-oriented radial model will be the same as those obtained from the input- and output-oriented models.

Tam (2004) also gave a non-oriented radial model operating under VRS, presented as (Eq. 1.8):

$$\begin{aligned}
 & \max_{g, \bar{\lambda}} && g \\
 & \text{subject to} && x_o - X\bar{\lambda} \geq 0 \\
 & && Y\bar{\lambda} \geq gy_o \\
 & && e_n\bar{\lambda} \geq 1 \\
 & && g \geq e_n\bar{\lambda} \\
 & && \bar{\lambda} \geq 0.
 \end{aligned} \tag{1.8}$$

In the case of VRS models, the efficiency scores from Eq. 1.8 will be less than or equal to scores for the same DMUs in both the input- and output-oriented VRS radial models, i.e. Eqs. 1.1 and 1.5. The multiplier forms of the non-oriented radial models can be found in Tam (2004).

## Additive Models

DEA is most useful for modelling production situations involving multiple inputs and multiple outputs. One of the inherent difficulties in dealing with these situations is the evaluation of trade-offs, for example between substituting one input for another. This evaluation is referred to as considering the mix or allocative efficiency of the DMUs. In situations with known prices for all inputs and outputs, the cost, revenue or profit can be optimized to decide upon the best input and/or output mixes. However, in many situations, prices or values are not known or not fixed for all inputs and outputs. Radial DEA models generally avoid dealing with mix issues by looking at proportional changes to inputs and outputs in their first stage. Proportional changes keep the input and output mixes the same as those originally employed by the DMU.

The additive model of DEA does address the input and output mixes of the DMUs. Its goal is to determine the maximum extent to which slacks can be removed from the DMU being evaluated. It is generally used as a non-oriented model, the VRS envelopment form of which is given as [refer to Cooper et al. (2007) – the multiplier form of the additive model can also be found therein]:

$$\begin{aligned}
 & \max_{\lambda, s^-, s^+} && z = e_ms^- + e_ss^+ \\
 & \text{subject to} && X\lambda + s^- = x_o \\
 & && Y\lambda - s^+ = y_o \\
 & && e_n\lambda = 1 \\
 & && \lambda \geq 0, s^- \geq 0, s^+ \geq 0.
 \end{aligned} \tag{1.9}$$

The characteristics of the additive model are very different from those of the radial DEA models. Its results are not easily expressed as standard efficiency scores,

i.e. values ranging from zero to one, with one representing efficiency. The optimal objective function value for efficient units in Eq. 1.9 is zero, as efficient units will have no slacks, and there is no defined upper limit on the total slacks. Unlike most forms of DEA (e.g. radial and slack-based measure models), the additive model can have zeros or negative values in the variable data, and is translation invariant, meaning that a constant could be added or subtracted from the values of a particular variable across all the DMUs without affecting the results. However, unlike most other DEA models, it is not unit invariant, and as such measuring a variable in miles as opposed to kilometers could affect the analysis results.

## SBM Model

Tone (2001) formulated the slack-based measure (SBM) as a development of the additive model that would generate a standard efficiency score and be unit invariant, while also allowing for input and/or output mix considerations. The envelopment form of the input-oriented CRS SBM is given by:

$$\begin{aligned}
 \min_{\lambda, s^-, s^+} \quad & \rho = 1 - \frac{1}{m} \sum_{i=1}^m s_i^- / x_{io} \\
 \text{subject to} \quad & x_0 = X\lambda + s^- \\
 & y_0 = Y\lambda - s^+ \\
 & \lambda \geq 0, s^- \geq 0, s^+ \geq 0.
 \end{aligned} \tag{1.10}$$

From Eq. 1.10, it can be seen that the SBM, like the additive model, is maximizing the total input slacks, but the slacks are considered as a proportion of the initial input value, as opposed to being considered in absolute terms. Similarly, output- and non-oriented, and VRS forms of the SBM, as well as corresponding multiplier forms of these models can be formulated, c.f. Tone (2001) and Cooper et al. (2007). The input- and output-oriented SBM models could undergo a second stage slack optimization, as occurs with radial DEA models, in the outputs and inputs respectively.

Comparing Eqs. 1.1 and 1.10, it can be seen that the SBM is similar in form and function to a radial DEA model. Whereas an input-oriented radial DEA model maximizes the proportional input contraction that is applied to all input variables, the SBM model maximizes the average proportional input contraction across all the inputs. Hence the SBM model is implicitly assuming that a 1% reduction in one input has the exact same value as 1% reduction in any other input, or as another example, the combination of a 0.4% reduction in a second input and a 0.6% reduction in a third input. Further, it can be noted that the efficiency score from an SBM model will be less than or equal to that from the corresponding (i.e. same orientation and RTS assumption) radial DEA model.