BENCHMARKING IN TOURISM AND HOSPITALITY INDUSTRIES

The Selection of Benchmarking Partners

Benchmarking in Tourism and Hospitality Industries

The Selection of Benchmarking Partners

KARL W. WÖBER

Institute for Tourism and Leisure Studies
Vienna University of Economics and Business Administration

CABI *Publishing*

CABI *Publishing* is a division of CAB *International*

CABI Publishing
CAB International
Wallingford
Oxon OX10 8DE
UK

Tel: +44 (0)1491 832111
Fax: +44 (0)1491 833508
Email: cabi@cabi.org
Web site: www.cabi-publishing.org

CABI Publishing
10E 40th Street
Suite 3203
New York, NY 10016
USA

Tel: +1 212 481 7018
Fax: +1 212 686 7993
Email: cabi-nao@cabi.org

A catalogue record for this book is available from the British Library, London, UK.

Library of Congress Cataloging-in-Publication Data

Wober, Karl W.
Benchmarking in tourism and hospitality industries: the selection of benchmarking partners / Karl W. Wober.
p. cm.
Includes bibliographical references (p.).
ISBN 0-85199-553-5
1. Tourism--Management. 2. Hospitality industry--Management. 3. Benchmarking (Management) I. Title.
G155.A1 W56 2002
338.4′791′068--dc21 2002000679

ISBN 0 85199 553 5

Typeset by AMA DataSet Ltd, UK.
Printed and bound in the UK by Cromwell Press, Trowbridge.

Contents

Preface viii
Acknowledgements xii

Chapter 1 Introduction 1
1.1 Development of Benchmarking 1
1.2 Types of Benchmarking 2
1.2.1 Process-based versus non-process-based benchmarking 4
1.3 A General Review of Benchmarking 5
1.3.1 The effectiveness of benchmarking 6
1.3.2 Concepts of benchmarking 7
1.3.3 Evaluating benchmarking theory 9

Chapter 2 Benchmarking Studies in the Tourism and Hospitality Industries 11
2.1 Principal Areas of Benchmarking in Tourism 12
2.1.1 Benchmarking in the hospitality sector 12
2.1.2 Destination benchmarking 15
2.2 IT Supported Benchmarking Systems 21

Chapter 3 Measuring Business Performance 23
3.1 The Input/Output Transformation Model 24
3.1.1 Multiple input–multiple output 25
3.1.2 Z-score analysis 26
3.1.3 The balanced scorecard 28
3.2 Types of Business Performance Evaluation 30
3.2.1 Definitions of competitiveness and performance 30
3.2.2 Concepts of comparative performance measurement 31
3.3 Important Terms Used in Performance Studies 36
3.3.1 Economical, allocative and technical efficiency 37

Chapter 4 Methods for Estimating the Production Function 42
4.1 Central Tendency Methods 42
4.1.1 Average ratio analysis 42
4.1.2 Simple regression models 43
4.1.3 Random coefficient models 46
4.1.4 Mixture models 49
4.2 Frontier Models 60
4.2.1 Data envelopment analysis 61
4.2.2 Stochastic frontier models 76

Chapter 5 The Austrian Hotel and Restaurant Panel 80
5.1 The Austrian Hotel Industry Between 1991 and 1997 81
5.2 Survey and Database Design 83
5.2.1 Limitations associated with accounting studies 86
5.3 The Basic AHRP Data Entry and Reporting System on the Web 88
5.3.1 Data entry 89
5.3.2 Data retrieval and analysis 91
5.3.3 The AHRP reporting functions 91
5.3.4 Financial key ratios in the AHRP 92
5.4 An Heuristic Procedure to Enhance the Basic AHRP System 92
5.4.1 The heuristic model 93

Chapter 6 A Mixture Regression Model 99
6.1 Sample of the AHRP 100
6.2 Input–Output Variables 102
6.3 General Procedure for the Selection of Comparison Partners 107
6.3.1 Specifications of the mixture regression panel data model 108
6.3.2 Identification of the optimal number of segments 109
6.3.3 Evaluation of individual companies 113
6.4 Limitations of Mixture Models for the Selection of Comparison Partners 118
6.4.1 General criticism of the central tendency method 119
6.4.2 Disregarding different characteristics of input variables 120
6.4.3 Problems with real-time benchmarking systems 121
6.4.4 Inadequacy of multiobjective problems 121

Chapter 7 Data Envelopment Analysis 123
7.1 Model Specifications 124
7.1.1 Analysis of individual years 125
7.1.2 Longitudinal analysis in DEA 129
7.1.3 Incorporating non-discretionary input factors 137
7.1.4 Procedure for ranking efficient companies 142

7.2 Strengths and Weaknesses of the DEA Approach 145
7.2.1 Stability of DEA results 146
7.2.2 Interactive DEA 148

Chapter 8 Evolutionary Data Envelopment Analysis 150
8.1 Motivation for an Evolutionary Data Envelopment Approach 150
8.1.1 Multi-modal characteristics of DEA problems 152
8.1.2 Optimization techniques in tourism research 155
8.1.3 Other AI approaches for the selection of optimal comparison partners 156
8.2 The Principal Idea of a Genetic Algorithm 157
8.2.1 GA approaches for constraint handling 158
8.2.2 GA extensions for multiobjective problems 159
8.3 Design of an EDEA 160
8.3.1 Data structures and main program 160
8.3.2 Constraint handling in EDEA 162
8.3.3 Initialization 165
8.3.4 The GA operators 168
8.4 Testing EDEA 181
8.4.1 Does EDEA find similar solutions compared to LP-DEA? 183
8.4.2 Does EDEA produce solutions faster than LP-DEA? 189
8.4.3 Is EDEA less sensitive to problem characteristics and parameters? 189
8.4.4 Is EDEA easier to implement than LP-DEAs? 190
8.4.5 Does EDEA have a higher applicability than LP-DEAs? 191

Chapter 9 Perspectives of Benchmarking Decision Support Systems in Tourism and Hospitality Management 193
9.1 Turning Experience into Breakthrough Results 193
9.1.1 Standardization 194
9.1.2 Measurement 195
9.1.3 Methodology 196
9.1.4 Information technology support 198

Appendix A 200
Mixture Regression Models 200
Data Envelopment Analysis 209

Appendix B 224
Evolutionary Data Envelopment Analysis (EDEA) 224

References 234
Index 257

Preface

Competition today is increasing fast, and customers are becoming more demanding every day. The tourism and hospitality industry is under as much pressure, if not more, than other industries, to manage and advance productivity. Benchmarking is one of the most recent management methodologies that has emerged for assessing the internal strength and weaknesses of a company and to evaluate the comparative advantages of leading competitors. The nature of benchmarking is about learning how to improve business activity, processes and management. The majority of managers today perceive benchmarks as a useful means to propose and implement performance improvements representing realistic goals based on other companies' actual achievements. However, despite its popularity in practice, benchmarking lacks a rigorous foundation in management science. Vaguely defined processes dominate the practice of benchmarking and most studies in this field are reviews of applications ('case studies') of only limited theoretical value.

One of the most crucial elements of the generic process of benchmarking is the selection of benchmarking partners. It is clear that an organization is unlikely to achieve effective results from its benchmarking initiative when it fails to select the right comparison partners. Although the identification of the optimal partners is arguably one of the most important factors for a successful implementation of benchmarking, very little attention is paid to methodological aspects related to this decision problem. Until 2000 no generally accepted methodology for the selection of benchmarking partners existed for easy adoption by most companies.

This book explains how to use various mathematical and statistical techniques in order to evaluate the best practice among the potential candidates for benchmarking. It reviews two different techniques for analysing the performance of companies and for selecting appropriate comparison partners. The objective of this book is to give methodological suggestions for tackling the three key questions faced in selecting the optimal comparison partner or the optimal set of comparison partners: (i) Who is the best practising company? (ii) Who can my company be compared with? and (iii) What are the goals which are most advantageous and achievable for my company?

This book provides a detailed description of a number of different performance measurement methods. In particular it compares central tendency and frontier methodologies, selected for their appealing characteristics in efficiency measurement. These methods differ in terms of the type of measures they produce and the assumptions that must be made regarding the structure of the data for their use to be valid. Through this book a reader should be able to understand the pros and cons of central tendency and frontier methodologies for performance measurement and the optimal selection of comparison partners.

Recently, data envelopment analysis (DEA), a non-parametric frontier methodology, has been suggested for identifying benchmarking partners. This book provides an introduction to DEA and its application to the present decision problem, and explains the most important extensions introduced in the literature.

The increased computational complexity necessary to overcome some of the limitations of more advanced DEA models induced the author to propose a completely new optimization algorithm for DEA problems. This approach involves an evolutionary computation technique which has been successfully applied to other optimization problems. The evolutionary DEA (EDEA) is basically an application and extension of the GENOCOP system, a floating-point genetic algorithm (GA) for constrained models developed by Michalewicz (1996). A significant proportion of the text is concerned with a discussion of advantages and drawbacks using standard linear programming versus evolutionary GA-based optimizers for solving DEA-problems. The EDEA approach provides a highly appropriate framework for building and solving DEA-models. This new model offers some striking advantages for the selection of comparison partners as well as for the ranking of efficient companies. Furthermore, the successive development of DEA-models along with flexible tools for model building and the increasing number of GA-based solvers are the basic arguments in favour of the use of EDEA rather than the traditional (linear programming-based) optimizer for DEA-problems.

Although the examples in this book are given in the context of hotel management, the approaches discussed could be used by any manufacturing or service-oriented firm. All the experiments in this study involve micro-level data. However, the methods considered can also be used for making performance comparisons at higher levels of aggregation. For example, one may wish to compare the performance of an industry over time or across geographical regions (destination benchmarking). The findings and analysis should be useful to an audience of scholars as well as practitioners involved in the measurement of performance. It also covers the relevant theoretical concepts in this field and should, therefore, be of interest to the academic community.

Throughout the book it is assumed that the person who is involved in conducting performance studies has a business database of potential comparison partners at his/her disposal. This assumption is not unrealistic as it reflects the increased importance of information technology and the developments of many inter- and intra-industry panel databases. Because it is not the author's intention to focus on the entire benchmarking process in detail, it was decided to choose 'The Selection of Benchmarking Partners' as a subtitle of the book.

Outline of Chapters

Chapter 1 serves as an introduction to the principal concept of benchmarking and why it became so popular within the industry. It describes various types and concepts of the benchmarking process which leads to a discussion on the theoretical foundations of benchmarking methodologies.

The second chapter consists of a comprehensive description of studies on the practice of benchmarking in the tourism and hospitality industry. Following the principal areas of benchmarking in the tourism industry, applications in the hospitality sector as well as in the destination management sector are thoroughly reviewed and the main limitations and shortcomings are discussed.

The third chapter deals with theoretical issues related to the measuring of business performance. It provides a theoretical framework for the measurement of business performance, gives an introduction to various techniques for its evaluation and, finally, explains important terms frequently used in performance studies.

Chapter 4 presents analytical techniques for estimating the unknown production function, basically classified as central tendency and frontier methods. For the former, the author particularly focuses on mixture modelling, a fairly new methodology that helps to overcome some of the parameter specifications necessary in ordinary regression approaches. For the latter, the author gives a detailed description of the merits and caveats of Data Envelopment Analysis, a recently very popular non-parametric approach.

Chapter 5 introduces the database used in the subsequent experimental studies and elaborates on existing performance studies in tourism research. Furthermore, it gives an example of a monitoring system which incorporates managerial judgements for the appropriate selection of benchmarking partners. Methods for performance measurement, which follow the classification into parametric and non-parametric techniques from Chapter 4, are applied and discussed in Chapters 6 and 7.

Chapter 6 focuses on the application of mixture regression models as a central tendency's approach for the selection of optimal comparison partners. The chapter includes an illustrative case example for the Austrian hotel and restaurant panel database and an in-depth discussion of the strengths and weaknesses of this method.

Chapter 7 provides an introduction to the mathematical approach to the estimation of frontier functions and the calculation of efficiency measures. It discusses the DEA models and illustrates these using numerical examples.

Chapter 8 starts with a discussion on the deterministic characteristics of DEA and the increased computational complexity necessary to overcome some of its limitations (e.g. treatment of outliers, fuzzy input/output specifications, dynamic data, ranking of efficient companies, etc.). The most significant limitations are caused by the properties of the linear programming technique usually applied in DEA. Therefore, in the course of Chapter 8, the author introduces EDEA, a completely new optimization technique for DEA problems to overcome at least some of the main restrictions. The new model offers striking advantages for the selection of comparison partners and opens new research questions for the future. Chapter 8 closes with a conclusion on the findings of the experimental study and a comparison with the linear-programming-based approach to the problem of optimal selection of benchmarking partners.

Finally, Chapter 9 discusses developments and issues that are relevant for the future of benchmarking decision support systems and summarizes the most important research questions surrounding them.

Acknowledgements

Writing a book is quite a project. Although I cannot guarantee this work to be free of errors of fact, logic and interpretation, I can assure the reader that there would have been many more such errors had it not been for the expertise, vigilance and kindness of the individuals who read and commented on all or parts of the manuscript or who took the time to discuss the book's topics with me. However, in acknowledging these many deserving individuals, I fear that I will inadvertently make some omissions. I can only offer my apologies if this turns out to be the case.

My first debt of gratitude is to Josef Mazanec for encouraging and helping me through the process of learning and writing this book. He was not only a constant source of insight and guidance on most of the topics in this book, but also an advisor, supporter and mentor throughout my professional career. Definitely, he will always be my professional benchmark.

I also want to express my gratitude to the other members of my tenure committee at VUEBA: Elmar Fürst, Stephan Klinger, Johannes Ledolter, Mikulas Luptáĉik, Gunther Maier, Petra Staufer-Steinocher, Alfred Taudes, A Min Tjoa, Klaus Weiermaier, Werner Weingraber and Hartmut Winner, for their critical comments which helped in revising the final draft of the manuscript.

I want to take the opportunity to thank Daniel Fesenmaier, not only for inviting me to become a research fellow at the National Laboratory of Tourism and eCommerce at the University of Illinois at Urbana-Champaign (UIUC), but also for providing countless hours of effort in discussing a number of issues described in the text. I also owe special thanks to Jiann-Min Jeng for his generous support during my time at UIUC and for introducing me to Asian culture. Special thanks are due to Jay Beaman who provided support, ideas and valuable feedback regarding materials in this text and to Karen Giunta for reading first drafts of the manuscript.

In preparing this manuscript I benefited from the support of many people at my home institution, the Vienna University of Economics and Business Administration (VUEBA). There, my special expression of gratitude goes to: Sara Dolnicar, Barbara Dutzler, Walter Ender, Klaus Grabler, Gregor Hoch, Margit Kastner, Nicole Mitsche, Ulrike Wandl and Andreas Zins. I particularly want to thank Ulrike Wandl who helped me many times with her patience and understanding. In addition, I would like to thank numerous friends for putting up with my strange behaviour during book crises.

I would like to recognize the considerable financial support from the Professional Hotel and Restaurant Associations within the Austrian Chamber of Commerce for running the Austrian Hotel and Restaurant Panel Study. My special expression of gratitude goes also to Rebecca Stubbs and the staff at CAB *International* for their support and help in the materialization of this book.

This book is dedicated to my whole family for their encouragement and support, in particular: Regina, Renate, Romana and Romana – the most significant four Rs in my life!

Karl W. Wöber
Vienna, October 2001

Chapter 1

Introduction

1.1 Development of Benchmarking

Benchmarking is a buzzword of the last decade of the 20th century. This management approach to identify 'who is best?' and 'what makes them so successful?' has experienced increased popularity, both in manufacturing and service companies. In management science, benchmarking is usually positioned as being an extension of an existing total quality programme, and as being a way in which to establish new, more relevant and efficient standards of performance. The increased interest in benchmarking has certainly been stimulated with the publication of Xerox manager Robert Camp's book on benchmarking (Camp, 1989). Since then, the phenomenon of benchmarking has been discussed by many authors, primarily in the form of management guidebooks (e.g. Spendolini, 1992; Watson, 1992, 1993; Zairi, 1992, 1996; Bogan and English, 1994; Karlof and Ostblom, 1994; Cook, 1995; Harrington and Harrington, 1996; Codling, 1998; Cross, 1998; Czarnecki, 1999). Benchmarking is about learning how to improve business activity, processes and management. Ahmed and Rafiq (1998: 228) point to the wide variation of commonly used definitions for benchmarking:

> Benchmarking is a continuous systematic process for evaluating the products, services and work of organizations that are recognized as representing best practices for the purpose of organizational improvement. (Spendolini, 1992)
>
> Benchmarking is a continuous search for, and application of, significantly better practices that lead to superior competitive performance. (Watson, 1993)
>
> Benchmarking is a disciplined process that begins with a thorough search to identify best-practice-organizations, continues with the careful study of one's own practices and performance, progresses through systematic site visits and interviews, and concludes with an analysis of results, development of recommendations and implementation. (Garvin, 1993)

> Benchmarking is an external focus on internal activities, functions, or operations in order to achieve continuous improvement. (McNair and Leibfried, 1992)

In summary, benchmarking can be defined as a systematic procedure of comparative measurement with the objective to achieve continuous improvement. The great success of benchmarking in recent years is probably related to its inherent characteristic of being a knowledge-sharing and motivational process. It encourages managers and their staff to think in terms of performance measures and practices to increase profitability.

Benchmarking is currently used in several business areas and particularly in quality management. The objective of benchmarking is the promotion of process or product improvement by the identification of a recognized standard and of the related actions required.

The insights gained from benchmarking provide an organization with a foundation for building operational plans to meet and surpass the standard and promote an overall awareness of business improvement opportunities. It is argued by practitioners as well as by scientific communities that benchmarking can promote thinking that generates improvement breakthroughs and leads to greater awareness of the need for long-term planning.

Despite the wide use of benchmarking techniques in quality, marketing, finance and technology innovation in the manufacturing industry, benchmarking is still a vague concept in the service industry, particularly in the tourism field.

1.2 Types of Benchmarking

According to the focus and methodology applied in a benchmarking endeavour, it is possible to distinguish four different types of benchmarking (see Table 1.1). A distinction made in the literature is the difference between external and internal benchmarking (Spendolini, 1992; Camp, 1995).

In 'internal benchmarking', units, branches, divisions or locations compare themselves with other units, etc. in the same organization. The evaluation is carried out by company management via questionnaires and audits and has the aim of improving the policy–targets–programmes–results loop. Internal benchmarking can also help company managers in identifying their strengths, weaknesses, opportunities and threats (SWOT) and therefore in improving economic efficiency of the company. Internal benchmarking frequently proves to be a useful learning experience and starting point for more adventurous benchmarking activities. Internal benchmarking is usually the first learning step along the path to external benchmarking; consequently, it is the most commonly practised form of benchmarking.

In external benchmarking, organizations compare themselves with other organizations, such as rival firms in the same industry ('competitive benchmarking'), non-competing firms in another industry ('best-in-class' or 'best-practice benchmarking') or aggregated data comprising a specific sector or

Table 1.1. Different types of benchmarking.

Approach	Scope	Objectives	Advantages	Disadvantages
Internal benchmarking	• Functions, departments, projects, businesses in the same company or group at the same or another location	• Improve competitiveness • Stimulate continuous improvement • Improve economic efficiency • Find effective employee rewarding systems	• Similar language, culture, mechanisms and systems • Ease of access to data • Existing communications • Relatively quick returns possible	• Might inhibit external focus and foster complacency • Possibly results in returns that are merely adequate
External benchmarking				
Best practice benchmarking	• Any organization regardless of sector or location	• Identify best management practices	• Possibility of breakthroughs • Broadens corporate perspective • Stimulates challenge • Less sensitive to ethical and political reservations	• Relatively difficult to access data • Change ramifications are greater • Higher profile
Competitive benchmarking	• Competitors (e.g. companies operating in the same sector)	• Identify performance, objectives, strategies and programmes of competitors • Identify best practices	• Similar structure and constraints • Relative ease of access to data • Relatively low threat • Helps to overcome complacency and arrogance	• Sector paradigms might restrain creativity • Legal, ethical and political considerations
Sector benchmarking	• Specific or similar sector or industry branch	• Identify sector strategies and programmes • Disseminate information on best practices • Define training packages	• Industry trends easier to assess • Relative ease of access to data	• More difficult to derive specific recommendations • Data also accessible to competitors

industry ('sector benchmarking'). In competitive benchmarking, processes are compared with those of competitors. This evaluation is mostly undertaken by strategic consultants using confidential information in addition to reports and other publicly available information.

The objective of best-in-class benchmarking is to identify best practice in a certain management area (e.g. marketing). Such an exercise frequently involves companies from different sectors distinguishing themselves for the effectiveness of their management systems. This kind of benchmarking exercise is frequently sponsored by a pool of companies willing to share information and suggestions for the improvement of their management system.

Finally, continuous improvement of business performance can be achieved by 'sector benchmarking'. Sector benchmarking is normally undertaken by industry associations with the aim of collaborating with authorities and the stimulation of business competitiveness by assessing the average performance of the sector and differences among individual companies.

Each of these options has distinct advantages and disadvantages. Table 1.1 summarizes scope, objectives, advantages and disadvantages of internal, best practice, competitive and sector benchmarking.

1.2.1 Process-based versus non-process-based benchmarking

Benchmarking is most effective when applied to processes which, when improved, would make a significant contribution to the business's overall competitive position. The decision regarding what to benchmark is therefore crucial for the overall success of a benchmarking project and requires careful examination of the core competencies and key business processes of a company.

Benchmarking often leads to the generation of qualitative and quantitative data. The two types of data are often referred to as practices and metrics (Camp, 1995). Practices are internal and external business behaviours which can cause the creation of performance gaps. The evaluation of organizational and managerial practices involves comparisons between processes and systems ('process-based benchmarking'). In this process-based benchmarking approach, success factors or critical factors are hypothesized by the benchmarking project team. This usually starts with purely descriptive activities where business processes are flow-diagrammed, effectiveness and efficiency measures are established, and interrelationships and dependencies with other processes are defined.

The measurement of critical success factors is the key driver for continuous improvement and the discipline required to eliminate complacency. Quantitative data provide a measure of performance at a particular time. Typical measures calculated and reviewed in a benchmarking study, are:

- financial performance indicators (business performance measures);

- technical performance indicators (productivity measures); and
- efficiency indicators (human contribution measures).

The compilation and comparison of quantitative data without particular focus on any business process can be described as 'non-process-based benchmarking'. Camp (1995) differentiates between process-based and results-based measures which is useful for better understanding the different concepts of process-based versus non-process-based benchmarking.

One common misunderstanding in benchmarking studies is that benchmarking refers to only one of the two approaches. Frequently, managers interpret the importance of focusing on processes rather than on outputs/metrics, listed in many benchmarking textbooks, as an indication that effective benchmarking can only be achieved by applying qualitative methodologies. However, it is a combination of both, by asking about their behaviour and practices (process-based) and process performance by identifying the extent of the gap (results-based), that benchmarking could be successfully operated. The difference, and examples of process-based and results-based measures, is illustrated in Table 1.2.

1.3 A General Review of Benchmarking

Although the benefits of benchmarking are readily perceived by every manager, they conceal different concepts and interpretations. Someone might argue it is 'old wine in new pipes' when he compares it with Porter's (1980, 1985) competitive analyses (e.g. Cross, 1998) or the goal-setting and goal-evaluation theory (e.g. Locke and Latham, 1990). Others believe it is a novel

Table 1.2. Differences and examples of process-based and results-based measures. (After Camp, 1995.)

Process-based measures	Results-based measures
• Relate to a particular process • Used by people working in the process • Maintained by people working in the process • Provide very quick feedback on performance • Visually displayed in the workplace	• Relate to broader issues or unit targets • Used more as management information • Data collected in workplace but analysed and presented elsewhere • Usually give retrospective results, often weekly or monthly • Often too detailed to be fully communicated but 'vital few' are displayed in workplace
Examples: • Effectiveness • Efficiency • Process consistency/variability • Quality output level	Examples: • Customer satisfaction • Employee satisfaction • Product performance • Financial performance

and promising management technique (e.g. Tucker *et al.*, 1987; McNair and Leibfried, 1992; Spendolini, 1992; Watson, 1992). The latter argue that, unlike competitive analyses, which stress the anonymity of contributors during the collection of relevant data, benchmarking is most successful when carried out as a partnership between two or more teams of people either from the same or from different organizations. Regardless of these varying perceptions of the objectives of benchmarking, most authors agree that, where competitive analyses help companies understand and improve their relative industry position, benchmarking enhances this and enables companies to learn from the best, regardless of sector or location (Lamla, 1995: 38).

1.3.1 The effectiveness of benchmarking

Derived from Camp (1989), Table 1.3 illustrates the difference in competitive behaviour with and without the practice of benchmarking. Companies who do not adopt benchmarking are usually characterized as:

- Internally focused, without a clear understanding of their strengths and weaknesses, a reactive approach to competitiveness and a poor knowledge of customers' true requirements. Feeble efforts to innovate are made.

Companies who do practise benchmarking can be described as:

- Proactive, externally focused and close to the markets they operate in. They have access to a limitless pool of ideas, use the market as a starting point for setting their objectives and have a very good understanding of customer requirements. They also tackle big problems to achieve quantum leaps in competitiveness.

Table 1.3. Reasons for benchmarking.

Objectives	Without benchmarking	With benchmarking
Change management	• Evolutionary change	• Ideas from proven practices
Industry best practices	• Few solutions • Frantic catch-up activity	• Many options • Superior performance
Defining customer requirements	• Based on history or good feeling	• Market reality
Establishing effective goals and objectives	• Lacking external focus	• Credible, arguable
Developing true measures of productivity	• Subjective, ignores market developments • Strengths and weaknesses not understood • Route of least resistance	• Solving real problems • Understanding outputs

Benchmarking therefore needs to be applied for the following reasons.

- It is an excellent strategic planning method, as it sets credible, easy-to-reach targets.
- It exposes organizations to state-of-the-art practices and, by instigating a continuous learning process, can help in the cultivation of a culture based on change and continuous improvement.
- It is an extremely good vehicle for education, involvement and empowerment of people and for optimizing their creative potential in the area of innovativeness.

Empirical studies testing the effectiveness of the concept and practice of benchmarking are rare. Recently, Mann *et al*. (1998) investigated the effectiveness of benchmarking in a field experiment carried out in an electrical products distribution company. The experiment involved 138 branches of a company, which were assigned randomly to a group performing benchmarking, and a control group. The dependent variable was percentage increase in sales performance over a 4-month period. Their findings suggest that comparison against partners and knowledge about 'best practice' contributes, together with 'goal-setting' and 'goal-evaluation', to the efficiency of benchmarking.

Similar findings were presented by Voss *et al*. (1997), who investigated the link between benchmarking and operational performance using a sample of over 600 European manufacturing sites. Again, in the study the authors could demonstrate that benchmarking may indeed contribute to improved operational performance, first through the firm's understanding of its competitive position and its strengths and weaknesses, and second through providing a systematic process for effecting change.

1.3.2 Concepts of benchmarking

The practice of benchmarking, as detailed by Camp and widely followed by practitioners, is dominated by the search for specific practices that will enhance performance with a controlled allocation of resources. This improved efficiency is achieved by the discovery of specific practices, typically for a single problem area, relying on simple engineering ratios. Typical benchmarking handbooks offer checklists for the conduction of a benchmarking project. These checklists are vaguely defined and sometimes even contradictory in process and content. For example, Camp's original ten-step benchmarking process (Camp, 1989) is described by Watson (1992) in the form of a six-step process (Fig. 1.1).

Codling (1998) synthesized different models that have been found to be successful benchmarking programmes. Her model comprises 12 steps arranged in four stages.

1. Planning
   - Select the subject area
   - Define the process
   - Identify potential partners
   - Identify data sources and select appropriate collection method
2. Analysis
   - Collect data and select partners
   - Determine the gap compared to benchmark
   - Establish process differences
   - Target future performance
3. Action
   - Communicate to management and others
   - Adjust goal and develop improvement plan
   - Implement
4. Review
   - Review progress and calibrate

Although stage 4 is a final review, constant monitoring and feedback should take place throughout the whole process.

A remarkable point here is that the identification of potential benchmarking partners is located in the planning stage of the benchmarking process, however, the final selection of one or more benchmarking partner(s)

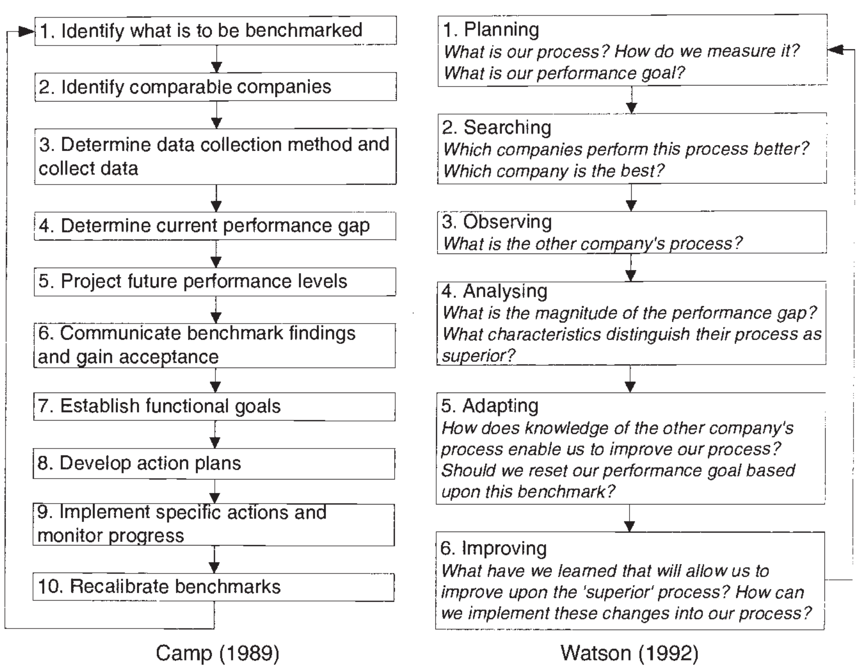


Fig. 1.1. Benchmarking process steps by Camp (1989) and Watson (1992).

takes place in the analysis stage once more detailed data have been collected. Similar to other authors, Codling does not provide any further information on what kind of procedures or methodologies can be applied to select the appropriate benchmarking partner.

1.3.3 Evaluating benchmarking theory

There is a vast amount of literature on benchmarking in management science; however, most of the studies are pure applications of only limited theoretical value. Very little attention is paid to methodological aspects in conjunction with benchmarking, especially to the right selection of benchmarking partners. Recently, Cross noted that 'one of the major weaknesses of many benchmarking studies is not spending enough time researching which companies might make relevant partners' (Cross, 1998: 9). However, the search for leadership companies and functions as introduced by Camp and others appears to be complex and lengthy. Bell and Morey (1994: 478) express this well, stating: '. . . the identification of leadership companies [in benchmarking studies] is as much art as science.'

Camp recommends the use of consultants, vendors and functional experts within the organization seeking help, as well as industry associations and public databases, as important sources of information for use in the selection of comparative companies. The level of detail provided for the selection of benchmarking partners seems insufficient, especially when compared to the efforts and costs involved in site visits and the implementation of change in the organization.

Recently, the application of four types of graphical techniques in evaluating benchmarking partners were discussed by Razmi *et al.* (2000). These graphical techniques, which the authors named 'alternatives–alternatives scorecard', 'shaded circles to portray scorecard-type result', 'ranking of alternatives' and 'polar graphs', are very simple forms of data visualization. In the generic form presented by the authors, they do not provide any guidance for the selection of benchmarking partners, as they completely ignore the problem of different priorities of attributes. To do this, the authors advise the use of more complicated methodologies, such as analytic hierarchy processes, knowledge-based systems and neural networks, but do not give any further information about how they should be implemented in a real-world application.

Many benchmarking efforts ignore, for the most part, differences in operating environments and service levels, and rely on simple engineering ratios. Such ratios are not appropriate when there are multiple outputs being produced with multiple types of resources. This is where traditional forms of multiple objective optimization techniques (e.g. linear combination of multiple attributes) fail. Very importantly, many of the resources being consumed, e.g. rent, utilities, labour and technology costs, are 'public' in nature in that they cannot be allocated to any output of the operation.

Further problems arise when simple benchmarking studies classify a data set of company statistics into a limited number of groups and provide means and medians for comparisons. The number and type of groups is mostly driven by expert judgements and is confined by the sample size and coding available in the data. It goes without saying that to be really useful, benchmarking partners should also match on levels of service provided, difficulty of operating environments, etc. Practices that are efficient and productive in one environment may not be relevant or helpful in a different environment.

In general one sees that more sophisticated initiatives spend more effort on the right selection of benchmarking partners. However, in the studies reviewed one does not find a system that could be easily adapted by other companies. The great majority of these studies are descriptions of counselling projects and do not leave a case-study-similar stage. This type of research is of very little value (and sometimes is even useless) for other companies who are searching for appropriate benchmarking partners. Therefore, it seems highly desirable that the methods for selecting leadership companies follow an explicit procedure which is valid for a broad range of companies. Furthermore, beyond isolating the most relevant benchmarking partners, there is also the need for methodologies which set up consumption targets and facilitate extensive sensitivity analyses.

The lack of systematic research in this area can be attributed to several factors. First, benchmarking is a relatively new phenomenon that only gained widespread attention in the early 1990s. Secondly, benchmarking is a practitioner-generated concept, and consequently, it is only loosely defined. This makes the concept difficult to study. Measures and tests for studying it need to be developed. Finally, benchmarking involves sensitive aspects of a firm's operations, in particular, information comparing them to their competitors. Thus, many firms are reluctant to allow access to independent researchers.

This text focuses on a general philosophy for identifying a given company's 'best practice' partners. Hence, it concentrates on tools facilitating step two of Camp's benchmarking process steps, i.e. building a peer group of relevant companies, operating in the same time period and matched on environment difficulties. It discusses the opportunities to measure performance gaps and introduces appropriate mathematical and statistical models that produce more meaningful results than those yielded by the comparison of simple descriptive measures such as mean scores.

Chapter 2

Benchmarking Studies in the Tourism and Hospitality Industries

Comparing performance figures is the procedure in the benchmarking processes that seems to be accepted and applied by tourism managers rather than more complex procedures like analysing or optimizing. Downie (1995, 1997) drew attention to the ongoing mismatch between the use and provision of information for planning and control activities in tourism businesses, pointing out the inconsistency of '. . . marketers working and planning with market segments, and accountants recording and reporting by operating department' (Downie, 1995: 214).

However, there is a growing body of research assuming that benchmarking is not solely a comparison activity. Studying the business performance of several hospitality sectors (attractions, restaurants and catering, motels), Bergin *et al.* (2000) found that benchmarking is often confused with the practice of competitive comparison studies. Terms such as benchmarking, interfirm comparisons and competitive comparison analysis are incorrectly interchanged. Benchmarking is considered to be a more powerful tool than competitive comparison analysis. It helps to keep a business focused on satisfying customers, improving procedures and achieving a world-class reputation, whereas competitive comparison analysis is merely the first data-gathering stage of the benchmarking process. Bergin *et al.* (2000) therefore propose that competitive comparison-analysis is only a component of benchmarking that lies in the final step of the benchmarking process of gap identification.

2.1 Principal Areas of Benchmarking in Tourism

A review of past benchmarking literature showed that there are a substantial number of both conceptual and empirical attempts to formulate a benchmarking approach. Focus and methodologies used in benchmarking studies in tourism can be very different according to the application field. In principle, benchmarking in tourism can be classified as follows.

1. Benchmarking of profit-oriented tourism businesses
   - Accommodation suppliers (hotels, motels, bed and breakfast places, pensions, camping sites, etc.)
   - Restaurants (all forms)
   - Tour operators and travel agencies
   - Airlines
   - Other profit-oriented tourism service providers (e.g. amusement parks, diving schools, etc.)
2. Benchmarking of non-profit-oriented tourism businesses/organizations
   - National or regional tourist boards/organizations
   - Attractions operated by public authorities or other forms of non-profit-oriented businesses (e.g. museums, galleries, theatres, operas, etc.)
3. Destination benchmarking
   - National benchmarking
   - Regional benchmarking
   - Local (rural or urban) benchmarking

The overwhelming number of benchmarking initiatives can be found among profit-oriented tourism businesses, particularly in the hospitality sector. Benchmarking in all other tourism areas has been very limited in terms of number of initiatives and in terms of their technical quality.

2.1.1 Benchmarking in the hospitality sector

Many researchers as well as practitioners have identified the importance of benchmarking for the hospitality industry. For example, Motwani *et al.* (1996) conduct a rigorous review of the literature relating to implementation issues of quality management in the hospitality industry, and show how the basic streams of definition and conceptual models relate to form the current practices of quality management within the industry. They propose a five-stage model for the implementation of quality management, and stress the need for benchmarking within this context. In another contribution, Kozak and Rimmington (1998) examine the role of benchmarking within the micro-structured hospitality sector. While benchmarking activity is growing in large organizations (e.g. Horwath International, 1998; Pannell Kerr Forster, 1998), there has been limited application among small hospitality businesses (e.g. Bottomley, 1995; Sundgaard *et al.*, 1998). Monkhouse (1995) examines the penetration of the small-to-medium-sized enterprise (SME) sector by the rapidly growing practice of benchmarking. Following a survey of over 200 SMEs, which clearly identified a ‘performance information gap’, the author has undertaken extensive quantitative and qualitative interviews with 25 senior managers. Findings provide a comprehensive picture of both current usage and the perceived or actual barriers to greater use of benchmarking. Monkhouse concludes that the practice of benchmarking in SMEs is embryonic and that little progress can be made by even enlightened managers until the barriers are understood. She also argues that a range of tools and techniques capable of accommodating the idiosyncrasies of small businesses need to be developed and made accessible. Additionally, Kozak and Rimmington (1998) find it significant that the examples of benchmarking carried out among small tourism businesses that they could find have almost all been carried out by external third parties, who first benefit from the data before they provide information back to the industry.

Still, there are benefits from tourism-related benchmarking studies. Breiter and Kline (1995) considered the role of benchmarking in hotel quality. Boger *et al.* (1999) identify and compare different levels of discounting among various lodging companies. The findings should assist managers in benchmarking the current discounting practices in lodging companies.

Based on a comprehensive exploratory study by Phillips (1996a,b), Phillips and Moutinho (1998a,b, 1999) propose a managerial tool, called the ‘strategic planning index’ (also ‘marketing planning index’), which measures the effectiveness of strategic planning (or marketing) activities and should facilitate a company’s benchmarking process. The tool was tested by a self-evaluation of 63 hotel managers in the UK hotel sector concerning three dimensions of performance: effectiveness, efficiency and adaptability. Effectiveness was measured by three ratios, namely occupancy percentage, average room rate and growth in sales per room; efficiency by return on investment and profit margin; adaptability by the number of successful new services/products introduced and the percentage of sales accounting for new services/products. A comprehensive list of strategic planning (or marketing) activities were factor analysed and studied by regression analysis to assess various performance indicators. The results of their study indicate that performance is an important measure of strategic planning and marketing effectiveness and that adaptability has a tremendous impact on the ability to maintain competitive advantages in operating efficiencies. In another paper generated from the same dataset, Phillips and Appiah-Adu (1998) focus on the value of benchmarking for the qualitative assessment of business processes. The authors review the concept of benchmarking and argue that the appreciation of a firm’s relative position is a vital component of strategic planning.

Min and Min (1996, 1997) chronicle the process by which a competitive benchmarking study of service quality provided by six luxury Korean hotels was carried out. In this study, attribute evaluations and weightings were determined by questionnaires completed by both employees and guests.

Although the authors generated benchmarking scores for each of the six hotels and provided sensitivity analysis for each of the service attributes, they did not give much emphasis to the selection process for benchmarking partners.

Dubé *et al.* (1999) conducted the most comprehensive study of the US lodging industry's best practices to date (September 2000). The study, which the authors themselves describe as a mammoth undertaking, was conducted under the umbrella of Cornell University's School of Hotel Administration and was financed by American Express.[1] The study resulted in a compilation of what the authors and a group of industry practitioners considered to be the most effective strategies and techniques used by the lodging industry's best operators. The selection of the best practice champions was based on managerial judgement. First the authors drew nominations for best-practice champions via a survey among 610 industry practitioners by mail, fax and e-mail, as well as from a website where individuals could download the survey. The survey resulted in 3528 nominations, including permitted self-nominations. Given this information the authors performed in-depth interviews with a list of 549 prescreened best-practice champions derived from the nominations received and the preparation of case summaries. From this information a total of 29 overall best-practice hotels was derived. In subsequent articles the authors have focused more closely on best practices in marketing (Siguaw and Enz, 1999a), food and beverage management (Siguaw and Enz, 1999b) and hotel operations (Siguaw and Enz, 1999c).

Without doubt, Dubé *et al.* (1999) spent considerable effort in the development of their study. However, their definition of 'best-practice' is purely judgemental as it lacks any hard-data evaluation. Particularly, to ask lodging industry managers to evaluate other hotels' performance assumes the managers have considerable knowledge and insights into the operation of their competitors. Therefore, it is not surprising that, during the final evaluation of their 'list of champion hotels' by frequent hotel customers and intermediaries, more than 50% of the overall champion's customers were unable to think of anything in particular that had created special value during their stay at a champion hotel. Despite this discouraging result, it also seems to be highly questionable if a single, purely demand-driven evaluation of the 'best-practice' companies is sufficient to validate their procedure. It is not unlikely, for instance, that a hotel which offers high value for low prices to customers, will have trouble in achieving its financial objectives in the long run.

[1] A focus section can be found in volume 40(5), and additional articles in volume 40(6), of the *Cornell Hotel and Restaurant Administration Quarterly*, summarizing the research findings on this project.

2.1.2 Destination benchmarking

Considering the tourism and hospitality industries, the majority of benchmarking studies can be found in the accommodation sector. Ritchie *et al.* developed operational measures for evaluating competitiveness and sustainability of a tourism destination (Ritchie and Crouch, 2000; Ritchie *et al.*, 2001). Although their framework provides a detailed set of measures and guidelines, their concept has not been empirically tested. Only a few attempts have been made to apply benchmarking methodologies for measuring the efficiency of regional tourism management or to assess the competitiveness of tourism destinations. Kozak and Rimmington (1999) review the literature on tourism destination competitiveness, stressing the requirement to establish which destinations are in direct competition. They note the importance of systematically evaluating such competitiveness both quantitatively through measurement of hard data (such as arrivals and tourism receipts) and qualitatively through soft data. To demonstrate the applicability of their proposal, the authors seek to benchmark Turkey as a destination against its competitors, which they have identified by means of a guest survey among British visitors. In their approach, destination performance is evaluated in various attributes (such as availability and quality of facilities, friendliness and natural environment) in comparison with other destinations, and complaints with regard to all destinations, in order to discover the relative performance of destinations.

Recently, the importance of benchmarking for the development of a systematic approach to tourism policy was also stressed by Alavi and Yasin (2000). They present a model based upon the 'shift-share' technique, benchmarking each country's growth in tourist arrivals from different parts of the world against that of the region as a whole. To illustrate their approach they perform market share analysis for various Middle Eastern countries and compare their performance with neighbouring European states.

Benchmarking information and communication applications for the purpose of marketing and sales in the tourism sector is a project conducted by the German Institute for Future Studies and Technology Assessment (IZT). The objective of this project, supported by the German Federal Ministry of Education and Research, is to develop and evaluate destination benchmarks (quantitative criteria) which allow the assessment of the strengths and weaknesses of tourism-related information technology (IT) investments (e.g. online information and reservation systems) in European destinations. The project provides information on best practices in website-presentations and suggestions on how tourism managers can improve the use of new media to meet international standards in a number of fields (see www.izt.de).

A case example of regional benchmarking demonstrated on a sample of 61 European cities was first presented by Wöber (1997). The original objective of the project, which was supported by European Cities' Tourism (www.europeancitiestourism.com), was to assess the volume and monitor the development of city tourism in Europe. For the first time, city tourism statistics have been regularly compiled and analysed leading to the most comprehensive database of urban tourism statistics. Data maintenance was organized online in close collaboration with the local city tourism representatives. First findings showed that most of the European cities enjoyed a significant upswing in their total number of bednights between 1986 and 1989. The economic recession that was first felt heavily in 1990 and the Gulf Crisis have led to a stagnation or decline of demand for many of the 61 cities from 1990 onwards. This unfavourable development in the beginning of the 1990s caused many tourism officials to start to investigate their position in more detail. In a first step, following an external benchmarking approach, European Cities' Tourism developed several reporting features, which allowed a comparison of the market segment performance of one city with other cities in Europe. This comparative tool was implemented in a Web-based decision support system and considered the heterogeneity inherent in international city tourism statistics by calculated median values for 29 different markets (countries of origin). An example for the kind of information which was suddenly available to the city tourism managers is illustrated in Fig. 2.1. It shows the development of Berlin's tourism measured by the number of bednights in all accommodations between 1999 and 2000. The relative figures for each generating market are then compared with the average development in all other destinations available in the database. Gains or losses in market share are identified and highlighted with different colours in the Web output.

As far as European urban tourism is concerned, performance studies frequently lead to a classification of cities into winner and loser categories. However, the classification cannot be based solely on their bednight statistics and other tourism targets also have to be considered. Examples of various success criteria, which are based on statistical sources and are easy to monitor on a regular basis, are given in Table 2.1. Some of the listed criteria are certainly correlated (e.g. high capacity utilization and growth in demand), however not absolutely. For an individual city manager, it might be a trade-off to enlarge its international volume by penetrating a single but very attractive market, and in the process disturb an equal guest mix configuration. A decision on the priority among the alternative strategies is a political one and will not be tackled here.

Table 2.1. Examples of success criteria.

Criterion	Possible measuring devices
Growth in demand	Growth rate in total number of bednights, arrivals or tourist receipts
Capacity utilization	Occupancy rate
Competitiveness	Market share
Internationality	Proportion of bednights from abroad
Seasonal distribution	Gini coefficient of monthly bednight statistics
Guest mix distribution	Gini coefficient of guest mix pattern

For measuring the smoothness of seasonal or guest mix patterns, the Gini coefficient has proved to be an ideal instrument. The Gini coefficient, which was initially applied to the seasonality interpretation by Wanhill (1980), is a bounded measure of inequality in the range from 0 to 1 and can be used to indicate the skewness of a distribution. The way to calculate the Gini coefficient is best illustrated by reference to the city of Lübeck (Fig. 2.2). Consideration of the degree of inequality begins with the joint cumulative distribution function of bednights, and the months in which they arise. First, the seasonal pattern for Lübeck has to be sorted and cumulated. If there is no seasonal effect the cumulative distribution function will be a straight line, the line of equality.

TourMIS - Netscape

File Edit View Go Communicator Help

	Berlin		All cities	
Market/origin	absolute	% p.y.	% p.y. (1)	Number (2)
Australia	45954	10.0	11.3	63
Austria	102489	10.6	-0.7	66
Belgium	73343	19.5	3.5	67
Canada	44250	31.3	16.5	67
Croatia			-2.0	28
Czech Rep.	20007	9.4		57
Denmark	176747	2.5	4.3	66
Finland	41384	-1.5	0.8	62
France	149123	13.0	3.8	71
Germany	8292587	19.8	3.4	70
Greece	30911	22.8	7.4	61
Hungary	15893	15.9	8.2	58
Ireland Rep.	13932	59.6	11.0	61
Italy	261014	37.1	-2.4	69
Japan	131786	20.0	8.4	67
Luxembourg	13253	44.1	18.0	57
Netherlands	204740	16.5	6.4	68
Norway	48694	7.6	11.8	66
Poland	48417	17.9	6.9	59
Portugal	23512	33.0	3.0	60
Russia	52796	22.9	9.9	60
Slovakia			-11.1	33
Slovenia			5.1	29
Spain	124077	27.5	5.9	68
Sweden	163115	11.5	6.9	66
Switzerland	172314	21.4	5.1	68
Great Britain	306941	23.9	9.9	69
United States	417452	32.9	10.5	69
Others	438194	35.3	7.6	78
Domestic	8292587	19.8	4.8	71
Foreign countries	3120338	22.2	6.5	73
All (domestic and foreign)	11412925	20.4	4.3	78

(1) = Mean change rate (median) compared to previous year.
(2) = Number of cities TourMIS could calculate a change rate.

Document: Done

Fig. 2.1. Destination benchmarking: comparing Berlin with other European cities (example taken from tourmis.wu-wien.ac.at).

The more unequal the seasonal distribution of bednights, the larger will be the area between the function and the line of equality. This is known as the concentration area and can be measured by an approximation formula named the Gini coefficient after its originator, Corrado Gini. The area is normalized to 1, thus giving the upper bound of the Gini coefficient. For complete inequality, the Gini coefficient is 1, and for complete equality it is zero.

For a permanent monitoring of multiple criteria, spider plots have proved to be an adequate graphical presentation (Harris and Mongiello, 2001). Figures 2.3–2.6 show five assessment criteria for selected European cities.

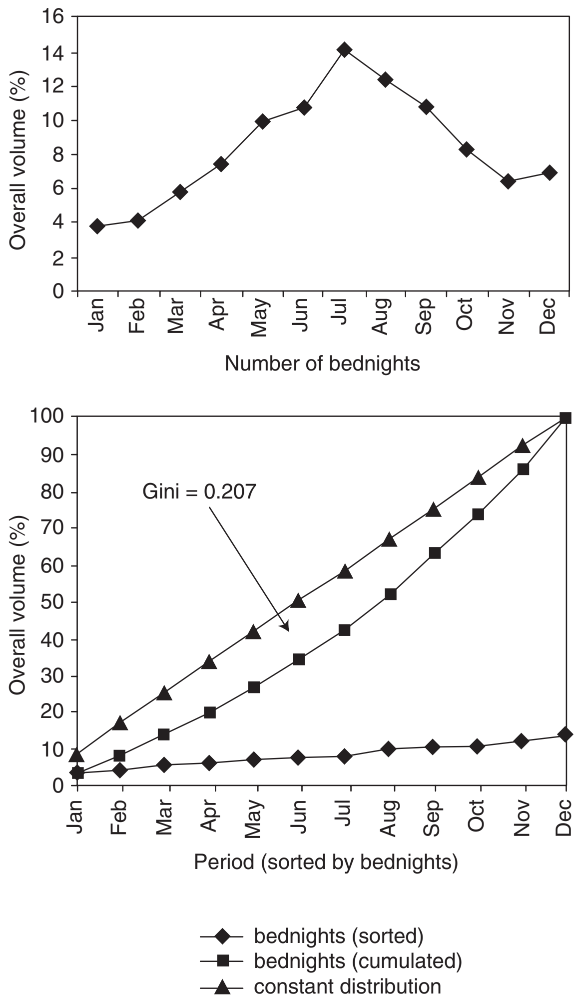


Fig. 2.2. Gini coefficient for measuring seasonality in Lübeck.

Each leg of the spider plot represents the value of a success criterion relative to competitors (39 cities could provide information for all items). The centre of the plot is the minimum value and the outer diameter represents the maximum value. The broken line represents the average number of all analysed cities and, when visible, indicates that the city's performance for this criterion is below average.

By displaying several plots simultaneously, differences and similarities among cities can be easily emphasized. Paris, which is obviously very successful in the European city tourism sector, still has some shortcomings concerning internationality, constant growth and equal guest mix distribution. Brussels and Amsterdam are certainly international metropolises with a similar problem concerning optimal capacity utilization. Heidelberg, which

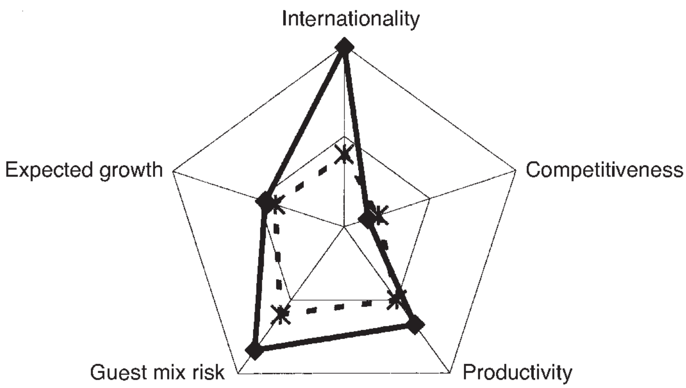


Fig. 2.3. Benchmarking destinations: spider plot for Amsterdam (◆). --*--, 39 cities.

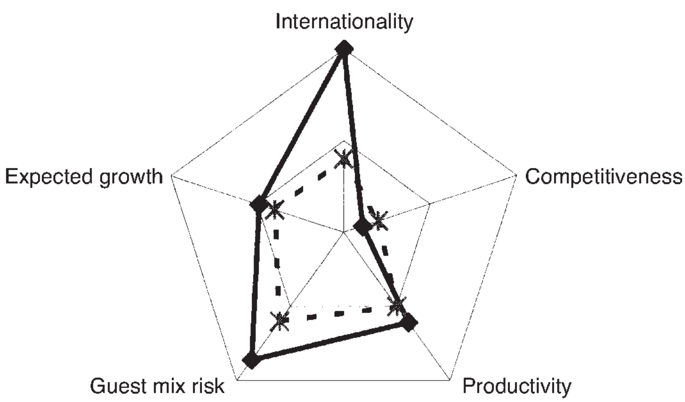


Fig. 2.4. Benchmarking destinations: spider plot for Brussels (◆). --*--, 39 cities.

shows a high occupancy rate in the evaluation period, is not very competitive and has an unfavourable guest mix distribution.

Wöber (1997) suggests a comprehensive indicator for the overall success of a city. The indicator is expressed by calculating the area defined by the lines of the spider's legs. The advantage of this kind of success evaluation, compared with an inspection of individual criteria, is that a high value can only be achieved if several assessment criteria show a favourable result simultaneously.

A hypothetical (probably non-existent!) optimal city leading all assessment criteria would find itself in the form of an equilateral pentagon. Normalizing by equating this area to 1 permits simple comparisons as percentage of optimal.

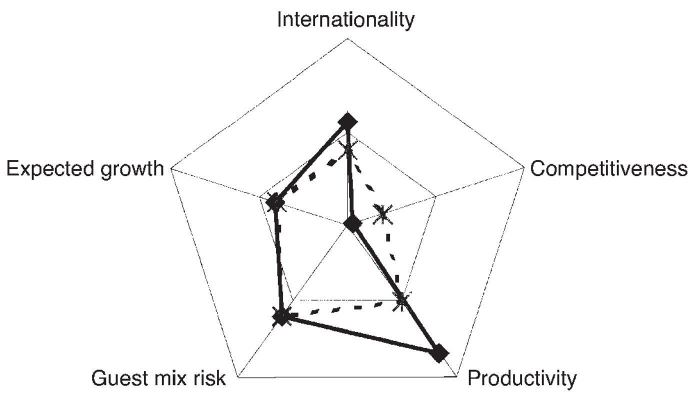


Fig. 2.5. Benchmarking destinations: spider plot for Heidelberg (◆). --*--, 39 cities.

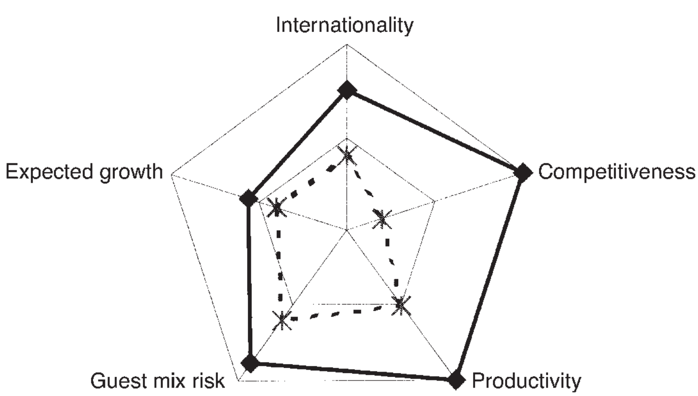


Fig. 2.6. Benchmarking destinations: spider plot for Paris (◆). --*--, 39 cities.

2.2 IT Supported Benchmarking Systems

A multidimensional performance measurement system, similar to the one described above, requires the gathering and dissemination of large amounts of information across many functions and levels of a tourism organization or destination. Considering the speed and flexibility of modern information and database systems it is obvious that this technology can be very useful for developing performance measurement systems (Brignall and Ballantine, 1996: 23; Halachmi, 2000).

Similar to the application in benchmarking in general, the hospitality industry has recognized the importance of technology earlier than other tourism sectors. There is a growing number of initiatives working on online systems in order to support the benchmarking process in the future.

The 'Environmental Benchmarking Tool' was developed by the International Hotels Environment Initiative and the World Wildlife Fund in the UK. Hotels can use benchmarkhotel.com to monitor their energy management, fresh water consumption, waste management, waste water quality, purchasing programmes, community relations and bio-diversity improvements. The management can compare their environmental performance with that of hotels with similar facilities in three major climate zones and design a program to reduce their costs and environmental impact. They can be entered directly into the database system where all individual hotel information remains confidential (see Fig. 2.7).

The HOST (Hotel Opportunity System Test) model developed by Econstat aims to reformulate the methodology of benchmarking for small and medium-sized enterprises (SMEs) and to adapt it to the tourism sector (Dall'Aglio, Vienna, 1999, paper presented at a Tourist Research Centre (TRC) meeting). Although the project still focuses on an Italian region (Emilia Romagna), it has a European scope. At the time of this research only a very simple prototype version of the program existed. In Austria, the Austrian Professional Hotel and Restaurant Associations, situated within the Federal Chamber of Commerce, have realized that Austrian SMEs cannot cope with this information deficit on their own. Therefore they decided to fund a research project which allows the industry to exchange data on business operations on a global basis in order to benchmark individual performances. The features of the system, which was developed on the World Wide Web, will be discussed in Chapter 5.

It is frequently argued that technology is a major force in providing competitive advantages especially in the areas of productivity, management decision-making, and education and training (Durocher and Niman, 1993; Go and Pine, 1995; Kluge, 1996; Kirk and Pine, 1998). However, the manner in which information systems are implemented generally evolves over time, and is rarely planned with the decision-making needs of executives in mind. In addition, the information systems used by managers often lag behind the techniques available to them. As a result, although an executive may be deluged with printouts, reports and statistics, he or she is not necessarily

Fig. 2.7. The environmental benchmarking tool (www.benchmarkhotel.com).

receiving the kinds of information needed to plan and to manage (Geller, 1985b; Umbreit and Eder, 1987).

Chapter 3

Measuring Business Performance

Since the objective of benchmarking is to improve a certain process, a benchmark must be 'better'. Although simple in concept, managers have found this difficult to achieve in reality (Codling, 1998: 184). Essentially, it means defining what is meant by 'better' in the context of the process in question as well as in the broader context of a company's operating environment.

Because of what needs to be achieved, the optimal selection of comparison partners requires a systematic study of a company's performance, thus a model of the underlying production process. Productivity is usually defined as the ratio of inputs and outputs. Input refers to the resources used in making a product or providing a service, while output is the product or service itself. Input and output may be measured in financial terms, i.e. costs and revenues, but not necessarily, as will be discussed later.

Several models have been introduced to measure performance in the hotel industry (Jones, 1988; Phillips, 1999; Southern, 1999). Jones (1988) presents a four-part model of service delivery based on inputs, intermediate output and outcomes. In order to describe this model, he gives a simple example based on the hospitality industry. A hotel's kitchen staff have available a range of commodities and foodstuffs, a level of expertise and a range of equipment which are all input factors to the production process. Each day the kitchen plans, prepares and serves a menu based on some assumption about the level of demand (intermediate output). But only a proportion of the dishes prepared are sold (actual output), and the hotel's customers that have purchased a meal have a wide range of experiences, such as dissipation of hunger, comfort, social contact, security, and so on (outcomes). Jones argues that productivity is largely determined by managing stage one, the relationship between inputs and intermediate output. Capacity is concerned in stage two, the relationship between intermediate and actual output; and quality is concerned with the issue of ensuring successful outcomes from actual output. Southern (1999) discusses the efficient use of business resources by reference to a generic operations management analysis framework. He finds that in the hospitality industry, management of performance, particularly in smaller and middle-sized enterprises, is ruled by intuition and past experience; he argues that a more systematic approach to process design, as practised by manufacturing and financial service companies, is needed.

Phillips (1999) speculates that competitive advantage can be achieved if several elements are congruent with business objectives. He suggests a model of inputs, processes, outputs, markets and environmental characteristics for building a performance measurement system in a hotel. Although the proposed performance framework is purely conceptual and does not involve any measurement system, it is a first attempt to build an input/output transformation model. A theory of measuring business performance therefore best starts with a closer look at some of the variables involved in this general input/output transformation model.

3.1 The Input/Output Transformation Model

The success of any business firm is a result of the interaction of two major sets of factors. The first major factors influencing the performance of a business enterprise emanate from inside the firm. They determine the firm's ability to use its resources to adapt to and take advantage of the constantly changing environment. Those inputs that are controlled or determined by the manager are referred to as controllable inputs to the model (Anderson *et al.*, 1997). Controllable inputs define the manager's decision alternatives and thus are also referred to as the decision variables or discretionary variables of the model (Fig. 3.1).

In any realistic situation, however, there may exist exogenously fixed or non-discretionary inputs that are beyond the control of a firm's management, and are therefore uncontrollable.[1] These uncontrollable variables are either factors determined by a company's market area (e.g. location of a hotel) or by physical characteristics of the property (e.g. number and mix of rooms). They are more or less exogenously fixed in the sense that they cannot be changed by management. In the flowchart in Fig. 3.1, the environmental factors are referred to as uncontrollable inputs to the model.[2]

There are several efficiency studies which have included uncontrollable variables. For example, Banker and Morey (1986b) illustrated the impact of exogenously determined inputs that are not controllable in an analysis of a

[1] The terms 'exogenously fixed', 'non-discretionary' and 'uncontrollable' variables, as well as the terms 'discretionary' and 'controllable' variables, are used interchangeably throughout this text.

[2] Although non-discretionary output is also conceivable, it is usually not addressed in performance studies.

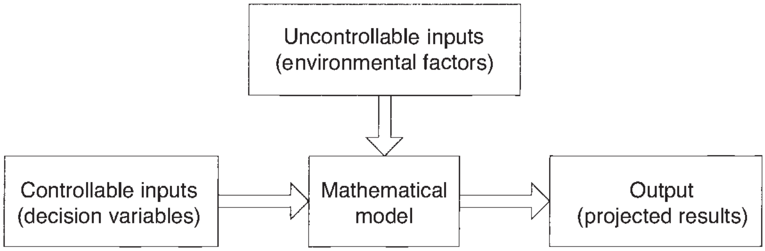


Fig. 3.1. Process of transforming inputs into outputs (Anderson *et al.*, 1997).

network of fast-food restaurants. In their study, each of the 60 restaurants in the fast-food chain consumed six inputs to produce three outputs. The three outputs corresponded to breakfast, lunch and dinner sales. Only two of the six inputs, expenditures for supplies and expenditures for labour, are defined as discretionary. The other four variables (age of store, advertising level, urban/rural location and drive-in capability) are beyond the control of the individual restaurant manager.

3.1.1 Multiple input–multiple output

As stated earlier, the definition of productivity is more complex than a single output measure. This is especially true in the public sector, where non-profit organizations (NPOs) provide social services particularly in health, education and defence. Profit is very rarely an objective. In the public sector in particular, the assessment of performance demands multiple objectives (e.g. Lewin and Morey, 1981; Lewin *et al.*, 1982; Sengupta and Sfeir, 1986; Smith and Mayston, 1987; Barrow and Wagstaff, 1989; Ganley and Cubbin, 1992; Valdmanis, 1992).

In the private sector, profit seems to be the dominant measure of output. However, the calculation of profit is hardly ever straightforward since it depends on sets of accounting conventions concerning the treatment of such factors as long-term investment, depreciation and tax-deferments (Norman and Stoker, 1991). Furthermore, only considering profit gives no indication of the potential for improvement within an organization or firm and therefore of their level of productivity. It is a common misunderstanding that information on profit is used to describe productivity. A profit number on its own conveys little information. It needs to be compared with, or put into the context of, some other number, measuring either a similar quantity in another organization (or the same quantity for another time period) or a related quantity in the same organization.

There are additional issues which have to be discussed when hospitality operations are considered. Firstly, there is a very large number and variety of inputs/outputs that occur in the daily operation of a hotel. Secondly, in the hotel industry, similar to the manufacturing industry, only the physical attributes of the hotel room or the food items might be considered standardized, whereas many of the other features experienced during the stay, such as service and atmosphere, are *intangible*. Thus, each service transaction with each individual customer can be regarded as unique. Given that both inputs and outputs include tangibles and intangibles, and these are almost impossible to measure directly, the realization of productivity improvement will almost always be difficult and the results are likely to be imprecise.

Another approach to the measurement is based on productivity objectives being a result of the various responsibilities a hotel has in terms of producing performance results. First, the management must achieve certain market performance results such as target sales volume, desired sales growth size or a competitive market share to maintain and strengthen the hotel's market position. Second, the owners (i.e. stockholders and creditors) expect the hotel to produce certain financial performance results in terms of profitability, growth and liquidity. Finally, a variety of other stakeholders in the business, such as employees, suppliers and the community, sometimes expect certain performance results in terms of employment stability and advancement, creditworthiness and good 'corporate citizenship' (for an overview of hotel stakeholders' interests see also Huckestein and Duboff, 1999). The degree to which a hotel meets the responsibilities can only be measured by several performance indicators simultaneously. Unfortunately, each individual number gives only a partial or incomplete picture and sometimes objectives are even contradictory (e.g. Pickworth, 1987).

Suppose a hotel has identified the following four key performance ratios for its operations:

- accommodation revenue per (whole time equivalent) staff;
- percentage of return visitors;
- overall guest satisfaction evaluated on a scale from 1 to 5;
- occupancy rate (capacity utilization ratio).

It would be operationally meaningless to simply add the four ratios to produce a composite overall measure. Some way needs to be found to accommodate all of these individual measures so that some sort of comprehensive assessment can be made. One approach is to weight each individual factor with the relative importance of the individual ratios, which became well-known as z-score analysis in the bankruptcy prediction research area.

3.1.2 Z-score analysis

In an attempt to reduce multiple measures into a single measure, some economists developed a viability indicator that has become known as the z-score (Altman, 1968). The z-score is a composite measure comprising the weighted sum of some of the key financial ratios. For example, a typical z-score might be computed as

$$z = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \tag{3.1}$$

where

$$x_1 = \frac{\text{profits before tax}}{\text{current liabilities}}, \quad x_2 = \frac{\text{current assets}}{\text{total liabilities}}, \quad x_3 = \frac{\text{current liabilities}}{\text{total assets}}$$

and b_0, b_1, b_2 and b_3 are constants.

In 1968, Altman introduced a bankruptcy classification model that applied discriminant analysis to two groups of companies over a period of time (Altman, 1968). The first group had either gone into receivership or voluntary liquidation, and the second group had remained solvent. The results of the z-score studies indicate potential significant application to credit-worthiness assessment and to external and internal performance analysis. Since this early work, there has been considerable interest in using quantitative models for bankruptcy classification, especially for credit-granting decisions. In fact, bankruptcy prediction has been a major research issue in accounting and finance since the early 1970s.

Most bankruptcy and related models are based on the concept of 'z-scoring' by use of weights usually determined as statistically significant coefficients of some linear statistical model, frequently the linear multiple discriminant model (Altman, 1968; Blum, 1974; Deakin, 1976b; Altman *et al.*, 1977; Sharma and Mahajan, 1980; Karels and Prakash, 1987; Messier and Hansen, 1988) and recently also neural network models (Wilson and Sharda, 1994; Wilson *et al.*, 1995; Lee *et al.*, 1996; Serrano-Cinca, 1996) and Data Envelopment Analysis (Barr *et al.*, 1994; Barr and Siems, 1997). Applications to the hotel industry have been reported by Olsen *et al.* (1983) who carried out a study on restaurant failure using univariate analysis. Using multiple discriminant analysis in the prediction of business failure in hospitality organizations was suggested by others (Adams, 1991, 1995; Adams and Kwansa, 1992).

This z-score approach is of interest because it attempts to give a comprehensive assessment of a company's viability that is comparable among a range of firms. However, there are several drawbacks in simple bankruptcy classification models. First, the dependent variable in such studies is whether a company went into liquidation or remained solvent and hence is defined as a discrete (qualitative, indicator) variable following a multinomial distribution. The statistical model chosen to represent the data must take this property into account. Linear discriminant models, however, implicitly assume that the attribute measurements arise from multivariate normal populations such that the classes have identical covariance matrices, differing only in the value of their mean vectors. More recently, artificial intelligence approaches to bankruptcy prediction models seem to overcome these statistical distribution assumptions (Odom and Sharda, 1990; Wilson and Sharda, 1994; Lee *et al.*, 1996; Serrano-Cinca, 1996). The main drawback of these measures is aimed at giving a single dimensional indication of the strength of a company – without regard for its standing with its competitors. Here the focus lies on the overall performance of an entity measured in comparison with the performance of several other entities.

Developments in the treatment of multiple objectives have also taken place in the broader context of performance assessment. Rusth and Lefever suggested:

> Some sort of multidimensional performance evaluation is much more appropriate in the international setting than the combination of net income and return on investment typically used for domestic operations. (Rusth and Lefever, 1988: 72)

This is consistent with the 'generic performance dimensions' proposed by Fitzgerald *et al*. (1991) on service businesses, the benchmarking methodology of Morey and Dittman (1995), the performance pyramid (Lynch and Cross, 1991), the integrated performance measurement proposed by Nanni *et al*. (1992) and with the work of Kaplan and Norton (1992, 1993) and Brander-Brown and McDonnell (1995) on the balanced scorecard. These emphasize the relevance of qualitative and quantitative approaches to performance measurement.

3.1.3 The balanced scorecard

One performance measurement method proposed to overcome the lack of 'balance' in performance measures is the 'balanced scorecard' introduced by Kaplan and Norton (1992, 1993). The balanced scorecard approach aims to provide management with a set of measures which combine to give a comprehensive view of the business. It is based on the idea that managers have to evaluate their business from at least four major perspectives: customers, internal business, innovation and learning, and financial (see Fig. 3.2). According to Kaplan and Norton, the performance measures developed to monitor these four perspectives should answer the following questions.

- How do customers view a firm?
- What business processes must the firm improve?
- Can the firm continue to learn and improve, and thereby create value?
- How does the firm appear to its shareholders?

The measures incorporated in the scorecard should provide a balance between external and internal measures, and thereby reveal the potential trade-offs between them. The balanced scorecard is intended to provide managers with a streamlined view of most major activities (Kaplan and Norton, 1997). The ability of the balanced scorecard to provide this view depends on the construction of a set of performance measures which will capture the pulse of a corporation in a few focused indicators. The implementation of a balanced scorecard requires that an organization has a clear view of where it is going, and how

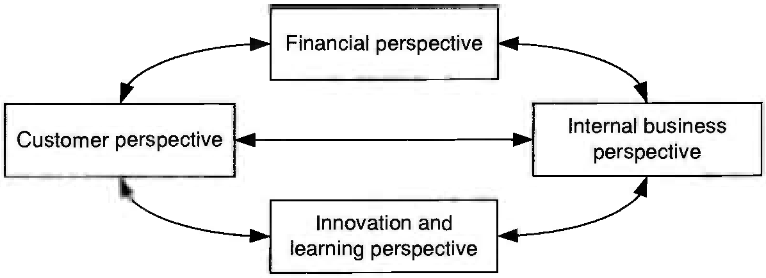


Fig. 3.2. Perspectives of the balanced scorecard (Kaplan and Norton, 1992).

activities, individuals' jobs and particular assets are linked to the overall objectives of the organization. The scorecard works via a process in which managers for each of the perspectives in Fig. 3.2 set goals, and specific measures for each are stipulated in order to achieve each goal. In this manner high-level goals are cascaded downwards into the organization through a process of tight specification while utilizing a consensus approach. In this way, the scorecard helps to translate and implement strategy.

Recently, Kaplan and Norton (2000) introduced the notion of the 'strategy map', showing how initiatives, resources and intangible assets will be converted into tangible outcomes. The 'balanced scorecard strategy map' is supposed to link the financial, customer, internal process, and learning and growth perspectives to the goal of improved shareholder value. The authors illustrate their approach with reference to Mobil North American Marketing and Refining, showing how the company moved from centrally controlled commodity product sales to become a decentralized, customer-driven organization (Kaplan and Norton, 2000).

The handbook publication of Kaplan and Norton's ideas on the balanced scorecard (Kaplan and Norton, 1996) has also created enormous interest in the hotel sector (Brown and McDonnell, 1995; Huckestein and Duboff, 1999; Schwärzler, 1999; Denton and White, 2000; Atkinson and Brown, 2001; Harris and Mongiello, 2001).

In general, the development of a balanced scorecard encourages the use of a broader set of measures. However, several problems associated with performance measurement remain unsolved. For instance, Gering and Rosmarin (2000) assert that although the balanced scorecard should empower decentralization, if badly implemented it can become a 'centralized trap' and part of corporate politics. The authors stress the importance of using the scorecard as 'the language of ongoing strategic discussion', driving it down to profit centre level and keeping the numbers of indicators manageable. They suggest that the learning and growth measures related to core competences are the only ones that can be dictated from the centre, and warn against incentivizing the scorecard directly.

Although Kaplan and Norton certainly spent considerable time and effort in the definition and selection of performance measures for a large number of companies, the problem of the relative importance of performance measures remains unsolved. Moreover, Kaplan and Norton do not review existing methodologies which could be considered for analysing these weights.

In general, it seems that the success of the balanced scorecard concept is attributable to the managers' wishes to overcome the information overload with which they are confronted in their daily operations rather than to have science-based findings. In this context, some recently published extensions to facilitate that are promising developments. For instance, Min and Min (1997) and Liberatore and Miller (1998) demonstrate the use of a multicriteria decision-making technique (Analytic Hierarchy Process) (AHP) and Larsen *et al*. (1997) propose an interactive simulation system incorporating expert judgements.

3.2 Types of Business Performance Evaluation

Determining how to measure the performance of a company is always difficult. First there is the problem of finding useful definitions of concepts such as competitiveness or performance (see, e.g. Buckley *et al*., 1988; Day and Wensley, 1988). Secondly, there is the problem of how to measure these concepts.

3.2.1 Definitions of competitiveness and performance

Studies regarding corporate performance have tended to use a variety of different measures of success, which can be classified into one of two groups: financial and non-financial. Researchers have employed financial measures such as profit (Saunders and Wong, 1985; Baker *et al*., 1988), turnover (Frazier and Howell, 1987), return on investment (Holley and Lynch, 1985), return on capital employed (Baker *et al*., 1988) and inventory turnover (Frazier and Howell, 1987).

Much of the criticism of traditional performance measurement systems stems from their failure to measure and monitor multiple dimensions of performance, by concentrating almost exclusively on financial measures. The use of non-financial performance measures is a relatively new area with growing recognition in the research community (Potter and Schmidgall, 1999). A number of recent studies suggest that non-financial performance measures such as customer satisfaction, internal quality indicators and improvement activities are related to future financial improvement (Saunders and Wong, 1985; Kaplan and Norton, 1992, 1993, 1996; Singleton-Green, 1993; Fisher, 1995; Bertels *et al*., 1999). In the hotel industry, Banker *et al*. (2000) analysed 6 years' worth of monthly data for 18 properties of a hotel chain. They documented that at this chain, non-financial measures related to customer satisfaction are related to future gross operating profit. They further report that the implementation of an incentive plan that incorporated non-financial measures resulted in an improvement of unit profitability. These remarkable findings have obviously not been well perceived by the industry. Recently, Mia and Patiar (2001) investigated the use of financial versus non-financial measures among 35 general managers of luxury hotels in Australia. Their results indicated that the general managers put far greater emphasis on financial indicators in evaluating the performance of their subordinates.

The choice of performance measures is one of the key challenges facing hospitality organizations today. Several initiatives have been made in order to identify and rank the criteria managers use to determine performance (e.g. Geller, 1985a; Icenogle *et al*., 1998). Although it is clear that no one performance measure will suffice, there is still uncertainty about what the attributes of good performance measures are and how many performance measures are needed for a firm to have a successful performance measurement system (Brignall and Ballantine, 1996).

One valuable attribute of a performance measure would be that the measure, whether financial or non-financial, is related to the value-generating process of a company, and thus is predictable (Potter and Schmidgall, 1999). Although this is an important attribute for the purpose of evaluation, another justification for the use of such a measure could be that it is helpful in guiding and directing employee behaviour by focusing the employees on activities that management wants the employees to attend to. If information about measures is to be communicated to employees, performance measures must be easy to understand, and, if used for determining compensation, they should be clearly related to employee effort and should not be easily manipulated by managers.

Geller (1985a,b) performed a study on the performance indicators or measures hotel executives deemed most important. His survey resulted in an ordered list led by occupancy percentage, average room rate, gross operating profit, rooms-department sales and rooms-department profit. When the survey participants were further asked whether their current executive information system supported them in meeting their responsibilities adequately, opinions were mixed. Half of the respondents said that it did and half that it did not. Those feeling that their systems were adequate mentioned the thoroughness and high level of detail the systems produced – particularly in the hotel operations and financial areas. The negative comments focused on the lack of marketing and competitive data of the systems. In summary, most executives felt that their systems provided adequate budgeting and historical information; many criticized the systems for their lack of timeliness and for failure to provide predictive and competitive data.

3.2.2 Concepts of comparative performance measurement

Industry 'rules of thumb' as benchmarks for business valuations are very commonly used by tourism managers. Mullen (1998) lists some examples from the hospitality industry and discusses the shortcomings of this kind of evaluation. The author concludes that such rules can only offer an early indication of expectations and are not a reliable measure of value. Especially in the area of profit statement analysis, evaluation has been enhanced by the growing recognition and use of more sophisticated comparative analysis techniques (Fay *et al*., 1976; Bernstein, 1978; Horwath *et al*., 1978; Monarcz and Portocarrero, 1986; Fenton *et al*., 1989; Harris and Hazzard, 1992; Atkinson *et al*., 1995; Coltman, 1998). Comparing financial information has proved to be effective for the assessment of internal operating performance related to budgeting and past results, whereas common size analysis facilitates operating performance comparisons on an intercompany and industry basis (Harris and Brown, 1998). Each of those methods encompasses the idea of a more informed approach to results analysis by the presence of 'absolute' and 'relative' measures of variation in performance. There are basically three ways that management can evaluate and control business performance.

1. Comparing budgeted versus actual input and output factors;
2. Comparing input and output factors over multiple periods; and
3. Comparing input and output factors with main competitors.

Each of these instruments has its strengths and weaknesses, which are reviewed in the following sections. It may be noted that the first two instruments can be implemented within a firm without incorporating any external data from other companies.

Budgeted versus actual performance

Performance is not absolute. For example, the success of a marketing campaign has a lot to do with top management's expectations versus results obtained (Bonoma, 1989). One formal way that management can express their expectations on business performance is through a budget. Operational planning and budgeting is familiar to most managers. A budget is centrally located at the core of a company's planning and control activities. The importance of the budgeting process for performance evaluation is stressed by many companies making the budget a clear target, which general managers and head office management are expected to meet. The budgets are, at least, intended to reflect a realistic performance target.

In the case of hotel groups, the keystone of the planning and control activities is the budgetary control system, which is an important part of the management accounting function (Kotas and Kreul, 1987). A survey regarding a significant number of hotels predominantly located in the UK showed that budgets in the hotel industry are usually prepared at a very detailed level over the main departments and the major overhead areas (Collier and Gregory, 1995). The preparation of the budget usually involves forecasts of the key elements such as occupancy percentage and average room rate, cost of sales, departmental employment costs and expenses and general overheads.

In terms of non-financial performance measures, dimensions identified were competitiveness, quality of service and resource utilization.

There are several methodologies relating to the use of budgetary information in a systematic performance evaluation of businesses. The most well-known techniques in the field of hotel and restaurant operations are the profit sensitivity (Kotas, 1978, 1982; Kotas and Wanhill, 1981) and profit variance analyses (Kotas and Kreul, 1987). The central concept behind the theory of profit sensitivity and profit variance analyses is that of the profit multiplier, which measures the impact on the net profit of a business with a given change in the relevant key factors (e.g. price level, sales volume, cost of goods sold, labour costs, etc.). By applying a very simple mathematical model it allows one to measure the effect of each key factor on net profit, thus allowing one to reveal the main reasons for profit deviations between the budgeted and the actual performance of a firm (in the so-called 'reconciliation' section of the profit variance analysis).

Longitudinal performance evaluation and control

When examining a company's operational performance over a period of time to introduce past performance into analysis, another element of comparison is used. A company that makes use of its resources more efficiently and generates greater revenues in the current period than previously is said to manage its operations successfully. In a simple two-period case the concept is very similar to the comparison of planned and achieved targets described in the previous section. Therefore, the methodologies suggested in the literature to analyse deviations in a two-period case are basically the same.

A concept of operational competitiveness solely based on estimating the efficiency of a hotel over several consecutive months was introduced by Parkan (1996). In his paper, a procedure to obtain a hotel's operational competitiveness profile, involving simple ratio-type computations that produce relative performance ratings, is proposed. There are also several studies that describe the merits of variance and regression analysis for the evaluation of business profitability over multiple periods of time (e.g. Monarcz and O'Brian, 1988; Russo, 1991; Harris, 1995).

Although there are a variety of statistical models that can be used to analyse cross-sectional, time-series data, the opportunity to compare performance profiles with other hotels was not considered by Parkan (1996). A comprehensive review of the statistical literature dealing with appropriate models is given by Dielman (1989). One of these models is the variance-components regression model that allows the error structure, but not the parameter estimates, to vary across cross-sections and time periods. Another model is the random coefficient regression model (Swamy, 1970) that generates a single set of response coefficients for all cross-sections by assuming individual cross-section coefficients are random deviations from a set of average coefficients. Leone *et al*. (1993) extend this random coefficient regression model to produce a separate set of coefficients for each cross-section.

Recently, a multi-period Data Envelopment Analysis (DEA) application for a small set of Australian hotel companies was introduced by Avkiran (1999). Avkiran used time-series data from secondary publications for a single hotel in a seasonal model (model no. 2 in his study) and data for seven hotels for demonstrating the usefulness of a window DEA[3] (model no. 3). Although Avkiran used an unrealistically small data set in his lodging industry case example, he was one of the first to dedicate a full textbook on DEA exclusively to the service sector (see also Westermann, 1999).

However, as Collier and Gregory (1995) point out, only a few hotel managers are applying any of these more sophisticated techniques. Their findings are also consistent with a study by Fitzgerald *et al*. (1991: 21,31). In general, hotel managers calculate variances for 2 consecutive years only on an item by item basis with the addition of percentages in some cases.

Competitive performance measurement

Intuitively it is clear that a key step towards achieving greater efficiency can be performance monitoring and measurement between 'organizational units' (Harrington and Harrington, 1995: 34). Organizational units may refer to internal departments, several outlets of one company or several companies within one industry. For example, by measuring the efficiency of its internal divisions, a company will understand its relative performance. This helps the managers to check if any appropriate corrective action needs to be taken and provides indications as to what kind of action, if any, is called for.

The complete model, which can be derived from the original input/output-transformation model (Fig. 3.1), is expressed in Fig. 3.3. Assume that there are n companies to be evaluated. Each company consumes varying components of m different inputs to produce s different outputs. Specifically, company j consumes amounts $X_j = \{x_{ij}\}$ of inputs ($i = 1, \ldots, m$) and produces amounts $Y_j = \{y_{rj}\}$ of outputs ($r = 1, \ldots, s$). For simplification reasons, the $s \times m$ matrix of output measures is denoted by Y, and the $m \times n$ matrix is denoted by X.[4] The measurement units of the different inputs and outputs need not be congruent and, in principle, smaller input and larger output amounts are preferable. Given the data, we are interested in measuring the efficiency of each company relative to all other companies in the data set by comparing the input–output relationships. In the absence of information about the true production function, a company j can be defined as efficient, when all of its input–output relationships are superior to its competitors (Equation 3.2) in at least the same or a more difficult environment (Equation 3.3).

[3] DEA window analysis is discussed in Chapter 7.

[4] In general, the text follows the notation used by Charnes *et al*. (1994b) in their major publication on Data Envelopment Analysis and by the majority of authors on business performance studies.

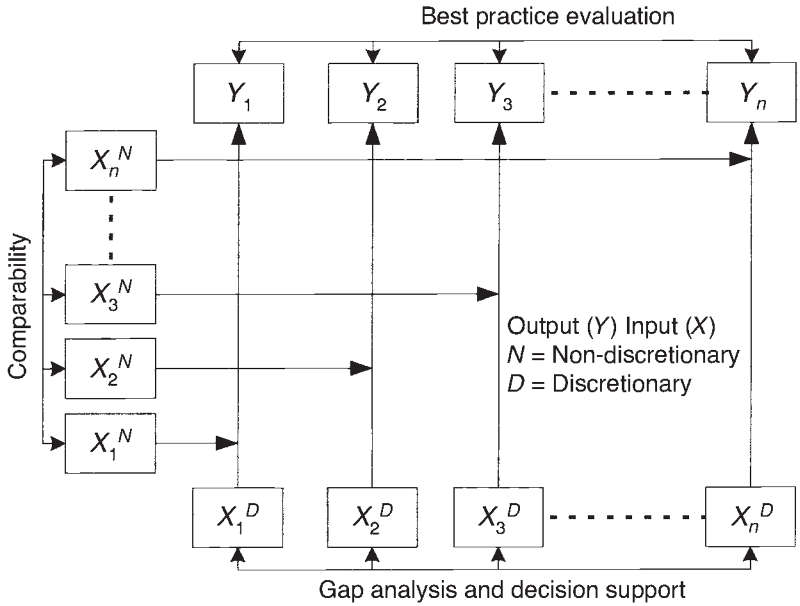


Fig. 3.3. The input/output transformation model in a competitive setting.

$$\frac{Y_j}{X_j^D} \geq \frac{Y_m}{X_m^D} \; (m = 1, \ldots, n) \tag{3.2}$$

$$X_j^N \leq X_m^N \; (m = 1, \ldots, n) \tag{3.3}$$

In reality, efficiency is not planned as such, since there is no known relationship between inputs and outputs, i.e. there is no known production function. Consequently, any assessment of efficiency is necessarily subjective, but competitive analysis is one possibility to overcome this problem.

The merits of competitive analysis were widely understood long before the publication of the presumably most significant competitiveness study undertaken by the Marketing Science Institute on the Profit Impact of Market Strategies in the 1970s (see Buzzell *et al.*, 1975). Through the widespread use of standard accounting systems and the digital processing and storage of comprehensive business data in database systems, it has become possible to determine industry benchmarks on a regional, national and international basis (Potter and Schmidgall, 1999). Rather than by solely referring to internally generated budgeted and previous performance results as in the past, the emergence of industry statistics databases has provided the facility for decision-makers to carry out operational analyses more effectively through the assessment of external norms and trends in planning and control activities.

The hotel industry is also catching up with the trend to use statistical databases, in the US faster than in Europe (probably caused by a higher proportion of SMEs in Europe compared with the US). This has primarily been led by key hospitality consultancy firms, which not only produce annual industry statistics for their own commercial purposes, but also directly benefit data contributors and others by making available valuable annual performance indicators and analyses of industry trends (Harris and Brown, 1998). Prominent leaders in this field are Pannell Kerr Forster Associates and BDO Hospitality Consulting, formerly Horwath & Horwath International. Both of these firms have been developing databases since the 1970s (Horwath International, 1998; Pannell Kerr Forster, 1998). In Germany, a similar database is maintained by the German Research Institute for Tourism (Deutsches Wissenschaftliche Institut für Tourismus (DWIF)), which publishes regular reports for Bavaria and Germany, covering a substantial sample of hotels and restaurants in the country. In Switzerland, a database of financial information on accommodation providers is available by the Société Suisse de Crédit Hôtelier (see www.sgh.ch). In Austria, the Austrian Society for Applied Research in Tourism (ASART) (see tourism.wu.edu/oegaf/oegaf.html) collects business data for hotels and restaurants on an annual basis and provides regular reports for the industry (Wöber, 1994–2000).

There are several problems that arise when comparing business data for several independently managed units. For instance, large companies may be able to obtain raw materials, credit, etc. on more favourable terms compared to smaller ones, enabling them to achieve lower costs. However, this does not reflect managerial efficiency since it is simply a company feature and not evidence of managerial skill. To achieve like-for-like cost comparisons one must control or eliminate the cost differences that are due to differences in company size. Having netted out the cost variation due to company size, the remaining cost differences across companies may still be due to differences in the organizations' objectives.

Another problem arises in conjunction with the reliability of comparing data from different sources. The level of precision and validity of external business data is sometimes unknown or subject to error. If any of the uncontrollable inputs in Fig. 3.3 are uncertain and subject to variation, the model is sometimes referred to as a stochastic or probabilistic model. The distinguishing feature of a stochastic model is that the value of the output cannot be determined even if the value of the controllable input is known because the specific values of the uncontrollable inputs are unknown. In this respect, stochastic models are often more difficult to analyse.

3.3 Important Terms Used in Performance Studies

In many papers written on the subject of performance measurement, there has been confusion and a lack of consistency in the use of terms like economical, allocative and technical efficiency. This section provides definitions and illustrative examples of these terms, which are frequently used in efficiency studies.

3.3.1 Economical, allocative and technical efficiency

Economists use the term productive efficiency to describe how well an organizational unit performs in utilizing resources to generate outputs or outcomes. In 1957, Farrell, a pioneer in this field, demonstrated that economical efficiency[5] can be decomposed into allocative efficiency and technical efficiency (Farrell, 1957).

A company is economically efficient when it produces a certain level of output at the lowest feasible cost. Costs may rise above the lowest possible level due to lack of either technical or allocative efficiency. Economic efficiency is a more inclusive requirement than technical or allocative efficiency, as both technical and allocative efficiency are required to achieve economic efficiency.

A company is technically efficient when it produces a certain level of output by using the minimum level of physical inputs. An example of technical inefficiency is when more people than necessary are used to carry out a certain task. A company is allocatively efficient when it uses inputs in the right proportion (for given input prices) to produce a certain level of output. An example of allocative inefficiency is when a high priced input is used when a cheaper one is all that is needed (i.e. wrong input mix), e.g. when company managers dedicate time to secretarial tasks such as typing (on a regular basis) instead of thinking how best to run the company (Cubbin and Tzanidakis, 1998).

The difference between technical and allocative efficiency is illustrated in Fig. 3.4, where it is assumed that output is produced by two factors x_1 and x_2, with the curve ψ being an output-isoquant. To bring allocative efficiency considerations into the picture, a budget (or cost) line associated with

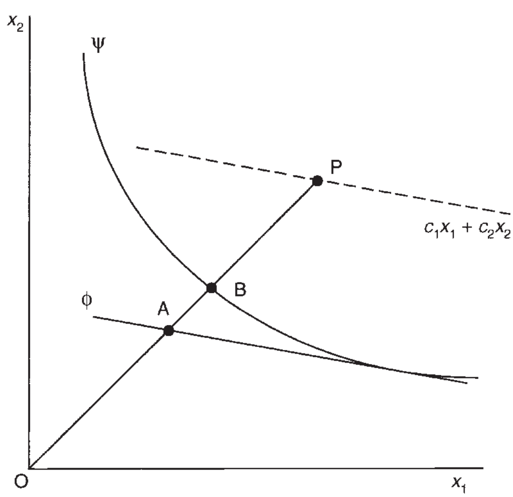


Fig. 3.4. Efficiency measures by Farrell (1957).

[5] Originally, Farrell (1957) introduced the term 'overall efficiency'; today the term 'economical efficiency' is more commonly used.

$c_1x_1 + c_2x_2 = k_1$, displayed by the broken line passing through P, is introduced. However, this cost can be reduced by moving this line in parallel fashion until it intersects with ψ. The new solid line ϕ represents the cost minimization plane with $c_1x_1^* + c_2x_2^* = k_0$, where $k_0 < k_1$ guarantees that total cost is reduced. The intersection between ψ and ϕ is the overall optimal point for company P as further parallel movement in a downward direction would be associated with reduced output. To measure company P's inefficiency let $\overline{OP}$, the line from O to P, cross ϕ at A and ψ at B. Given this, the economical efficiency (sometimes also referred to as 'overall efficiency') of unit P is measured by:

$$E = \frac{OA}{OP} \tag{3.4}$$

Technical efficiency (T), measured as the radial distance that P is from the isoquant, and allocative efficiency (A), measured as the radial distance from the cost minimization plane, are given by:

$$T = \frac{OB}{OP} \text{ and } A = \frac{OA}{OB} \tag{3.5}$$

Note that allocative efficiency provides a measure of the extent to which the technically efficient point, B, falls short of achieving minimal cost because of company P's failure to make the reallocations necessary to move from B to the overall optimal point, which is located at the intersection between ψ and ϕ.

Economical efficiency (E) can also be computed from A and T as follows:

$$E = \frac{OA}{OP} = \frac{OA}{OB} \times \frac{OB}{OP} = A \times T \tag{3.6}$$

In summary, firms can operate suboptimally for two fundamental reasons. The first is the failure to allocate resources in the most efficient manner (allocative inefficiency). The second way is related to a firm's ability to utilize its resources given their allocation or technical inefficiency. In other words, two firms may have exactly the same resource allocation, yet one firm produces less output than the other. The difference between how a firm could potentially utilize its resources versus its actual utilization is termed 'X-inefficiency' (Leibenstein, 1966; Anderson *et al.*, 1999).

The majority of X-efficiency losses, according to Leibenstein (1966), arise from inadequate motivation by firm management. He also suggests that motivation levels are linked to the structure and competitiveness of the market in which a firm operates. If managers and/or workers could be encouraged or persuaded to work more effectively, firms would improve performance without changing their resource allocation. If a firm is operating in a competitive market, managers and workers may feel pressure to work more efficiently and vice versa. In other words, there is another relationship that must be considered because obtaining a high-efficiency estimate is more likely to be observed in a competitive market and obtaining a low-efficiency measure is consistent with a less competitive market. Thus, care should be taken in making definitive statements, as efficiency measures do not prove market or firm efficiency.

There are two empirical approaches to the measurement of efficiency based on the above concepts of technical and allocative efficiency. The first, favoured by most economists, is parametric (either stochastic or deterministic). Here, the form of the production function (the isoquant ψ in Fig. 3.4) is either assumed to be known or is estimated statistically. The advantages of this approach are that any hypotheses can be tested with statistical rigour and that relationships between inputs and outputs follow known functional forms.

However, in many cases there is no known functional form for the production function and, in some cases it may even be inappropriate to talk in terms of such a concept. This becomes clear when someone considers the case in public sector organizational units that are not, for example, concerned with taking unfinished goods (or raw materials), processing them and producing finished goods for sale or transfer.

In the parametric approach the functional form usually chosen is Cobb–Douglas. In this context the Cobb–Douglas functions are estimated by 'averaging' statistical techniques, such as regression. Each unit is then compared with an average, but it is not immediately clear what this average represents. It clearly does not refer to a firm of 'average size' nor indeed to a firm having 'average means at its disposal' (or 'average technology').

In the non-parametric approach no assumptions are made about the form of the production function. Instead, a best-practice function is built empirically from observed inputs and outputs. This will necessarily be piecewise linear and, as such, is an approximation of the 'true' function, if one exists. In this case, the observed points are assumed to provide empirical evidence that production is possible at the rates specified by the coordinates of any point in this region (Cooper *et al.*, 2000: 7). For example, Fig. 3.5 shows observations for a number of similar companies, P_1–P_{10}, where the axes are input per unit output produced. From the efficiency point of view, it is natural to judge companies that use smaller inputs to get one unit output as more efficient. Therefore, companies P_8, P_2, P_7 and P_{10} are identified as efficient, as there is no other company that produces the same amount of output with less input. The line joining P_8 to P_2, P_2 to P_7, and P_7 to P_{10}, designates the efficiency 'frontier', which is assumed to extend parallel to the axes beyond P_8 and P_{10}. Technical, allocative and economical ('overall') efficiencies are calculated in an analogous manner to the approach shown in Fig. 3.4.

The inefficiency of companies not on the frontier line can be measured by referring to the companies that build the frontier. For example, company P_4 is inefficient. The technical inefficiency of this company is represented by the line from zero to P_4 divided by the line from zero to B. Hence, the inefficiency of P_4 is to be evaluated by a combination of P_2 and P_7 because the point B is on the line connecting these two points. Companies like P_2 and P_7 are commonly referred to as the 'reference set' or 'peer members' for company P_4. The reference set for another inefficient company, P_3, consists of companies P_2 and P_8, which clearly shows that the reference set may differ among inefficient companies.

Similar to the example illustrated in Fig. 3.4, economical efficiency can be calculated. Given input prices, the isocost line is reflected by the broken line passing through P_4. Reducing the total costs by moving this line in parallel fashion until it intersects with the frontier at P_7 gives A, the point that determines the economical (or overall) inefficiency of P_4. The relative distance between zero and A and zero and B measures the amount of allocative inefficiency of company P_4. In Fig. 3.5 only company P_7 is both technically efficient and allocatively efficient, whereas companies P_8, P_2 and P_{10} are technically but not allocatively efficient.

From the simplistic case shown in Fig. 3.6, it can be deduced that company P_1 is the most efficient and, if no other management units are included in the

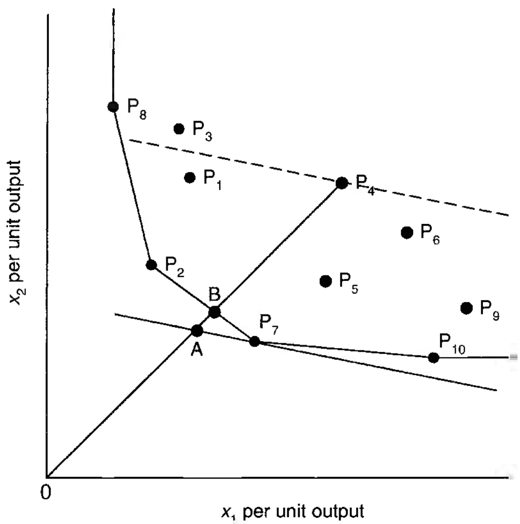


Fig. 3.5. A piecewise-linear efficiency frontier.

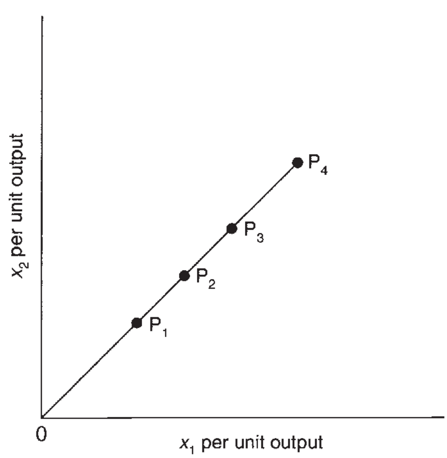


Fig. 3.6. Comparative efficiencies.

analysis, P_1 can receive a reference efficiency score and scores for P_2, P_3 and P_4 relative to P_1 can be computed. Thus, if P_1 is assigned an efficiency score of 1, it is possible to say that P_1 'is efficient relative to P_2, P_3 and P_4'. Therefore:

$$1=\frac{OP_1}{OP_1}>\frac{OP_1}{OP_2}>\frac{OP_1}{OP_3}>\frac{OP_1}{OP_4} \tag{3.7}$$

and the ratios for P_2, P_3 and P_4 determine their own relative efficiency scores. This is analogous to a Leontief single process input–output system. Farrell extended it to cover many processes and many inputs.

Chapter 4

Methods for Estimating the Production Function

4.1 Central Tendency Methods

4.1.1 Average ratio analysis

Thanassoulis *et al*. (1996) refer to the ratio of an output to an input, or an input to an output, as a performance indicator. Ratio analysis typically involves the use of a number of performance indicators and has traditionally been the method of choice in assessments of performance. In single input–single output contexts a performance indicator is a meaningful, easy to use measure of performance. However, this is not the case when multiple non-commensurate inputs and/or outputs are involved. The difficulty, highlighted also by Barrow and Wagstaff (1989) and Greenberg and Nunamaker (1987) among others, stems from the fact that each performance indicator reflects only one input and one output level and so it is difficult to gain an overall view of the performance of a company when not all its performance indicators indicate a similar level of performance.

Performance ratios are widely used throughout all sectors of business and commerce. The best-known ratios are for financial and production management, but ratios have also been developed to assess marketing, purchasing and personnel management. Even in such areas as the accountancy and consulting professions, use is made of measures of performance.

There are two different types of ratios that have to be distinguished in performance measurement. The first type refers to ratios which describe the relationship between two input or two output variables. In general, these ratios are used to better understand structural differences among companies rather than to measure performance. However, they are crucial during the search for appropriate benchmarking partners when comparability of business characteristics plays an important role. An example of this type of ratio is the 'acid test' (quick assets × current liabilities^{-1}), which indicates how a company would react if it were called upon to settle its current liabilities.

The second type refers to ratios of outputs to inputs. Larger values of such ratios are associated with better performance. This special form of financial ratio is called a productivity ratio (Coelli *et al.*, 1998). The issue of productivity should be at the heart of any decision-making organization. Whether the prime aim is higher profits, higher overall shareholder value and/or larger market share, or more satisfied customers, companies operating in both competitive and regulated markets need to improve productivity if they are to outperform their competitors. Management science usually distinguishes between total factor productivity, which is a productivity measure involving all factors of production within a company, and several partial measures of productivity, e.g. labour productivity or capital productivity (profitability). Organizations use a number of productivity measures to assess their performance. One of the most powerful financial productivity measures is the 'return on investment' ratio. This is used particularly when it is proposed to build or purchase a processing plant, shop or hotel, or to make another addition to the company's assets.

When the production process involves a single input and a single output, the calculation of productivity measures is a trivial matter. However, when there is more than one input (which is usually the case) and/or more than one output, then more sophisticated techniques for the construction of productivity measures must be applied.

4.1.2 Simple regression models

In the context of cross-sectional analysis, various authors in the business performance research field consider the case where the ratio of two accounting variables Y and X is compared to some characteristic value, b. If Y is proportional to X, then for the ith company the difference between $Y_i X_i^{-1}$ and b can be interpreted as an effect attributable to the individual company, i.e. as an indication of that particular company's departure from the norm. Thus, under the assumption of proportionality, inferences may be drawn directly from financial ratios (McDonald and Morris, 1984, 1985). But Barnes (1982) and Cubbin and Tzanidakis (1998) have suggested that regression analysis may be preferable to simple ratio analysis.

Consider a sample of companies with varying cost/output combinations, denoted as $\{X, Y\}$. For a given level of output Y, the observed operational costs X of each company can be thought of as the costs of an average-efficiency company, plus an efficiency factor u reflecting the efficiency difference ('costs inefficiency') between the particular company and the average-efficiency company, simply formulated as

$$X = \frac{1}{n}\sum\{X, Y\} + u \tag{4.1}$$

To measure relative efficiency of a particular company, an estimate or benchmark of the costs of the average-efficiency company with the same level of output as the company in question is needed. This can be obtained with the help of regression analysis.

Figure 4.1 illustrates observations and a regression through them. Operating costs X are plotted on the horizontal axis and output Y is plotted on the vertical axis. The dots indicate results for companies with certain cost–output combinations. The fitted regression line provides an estimate X^* of the average efficiency costs for a given level of output Y. For example, in the case of company P with observed costs X_P and output Y_P the estimated average efficiency cost (benchmark) is given by X_P^* which corresponds to the point where the dotted line intersects the estimated regression line. The difference between the observed and the estimated benchmark costs ($X_P - X_P^*$) is the estimate u_P^* of efficiency u_P. Similarly, for each company, an estimate u^* of its true efficiency u is given by the difference between the observed and estimated average efficiency costs $u^* = X - X^*$. Companies below the regression line are of below-average efficiency whereas companies above the regression line are of above-average efficiency. A scale independent efficiency score for each company can be calculated by expressing the difference between observed and predicted costs as a percentage of the predicted costs, i.e. the efficiency score is given by

$$e = \frac{X - X^*}{X^*} \tag{4.2}$$

The regression analysis assumes that the average efficiency company increases linearly with output Y:

$$Y = \alpha + \beta X + u \tag{4.3}$$

According to the above, a unit increase in output increases costs by β, and for zero output there are fixed costs α.[1]

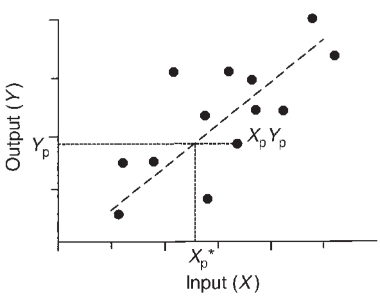


Fig. 4.1. Regression line and points fitted in an analysis for performance measurement.

[1] In an output-oriented model, α will be negative.

The linear regression model based on the ordinary least squares (OLS) algorithm to make estimates involves some statistical assumptions. This set of statistical assumptions enables, in addition to measurement, the proper statistical estimation and hypothesis testing of the parameters used in the model. The OLS algorithm fits the line by minimizing the sum of squared deviations of the observed costs from the line. Because regression analysis provides conditionally, upon each level of output, an estimate of the average efficiency costs, it is often said to fit an 'average line' to the data, or more accurately, a 'conditionally average line'.

Log transformation

Many accounting variables are sums of similar transactions with constant sign, for example: sales, stocks, creditors and current assets. Unlike other variables, such as earnings, cash flow and net working capital, which can have both positive and negative values, such accounting variables are bounded at zero. Therefore, for large samples of companies, the evidence is that the distribution of the first group of variables may be Pareto-like, or log-normal (Deakin, 1976a; Ijiri and Simon, 1977). Furthermore, a ratio of two such variables has the same properties (Lev and Sunder, 1979). That is, if X and Y are distributed log-normally then, due to the additive properties of normal variables, $\log Y - \log X$ is distributed normally, hence the ratio YX^{-1} is also log-normally distributed. Therefore, regression studies in the business performance field, especially when dealing with accounting ratios, adopt the use of log-normal regression analysis.

For example, in the case of a log-normal transformation, Equation 4.3 can be redefined as

$$\ln Y = \alpha + \beta \ln X \tag{4.4}$$

For $\beta = 1$, this is

$$\ln Y - \ln X = \alpha \text{ and, therefore, } \frac{Y}{X} = \exp(\alpha) \tag{4.5}$$

where α is the mean of the logarithms of ratio YX^{-1}. Thus, $\exp(\alpha)$ can be referred to as a 'benchmark', or characteristic value against which the ratio of the *i*th company $Y_iX_i^{-1}$ can be compared. The OLS estimate of α is the average of the logarithms of the ratios

$$\alpha = \frac{1}{n}\sum \ln\left(\frac{Y_i}{X_i}\right) \tag{4.6}$$

and so an estimate of the characteristic value $\exp(\alpha)$ is obtained as the geometric mean of the ratios $Y_iX_i^{-1}$. For a bivariate log-normal regression, the model can be summarized as

$$\ln Y_i = \alpha + \beta \ln X_i + \varepsilon_i \tag{4.7}$$

where unexplained variability is represented by ε_i, which is assumed to be uncorrelated and normally distributed with mean 0, variance σ^2.

Here it is worth noting that, for log-normal variables, the geometric mean is the median. In the case of financial ratios of accounting variables, which are sums of similar transactions, this estimate should be preferred to its competitors like the arithmetic average of the ratios or the ratio of arithmetic averages, which implicitly rely on the normal distribution of the ratios or their components. Elsewhere, it has been shown that, in cross-section, the log-normal model can provide a useful approximation for the ratio of accounting sums (McLeay, 1986).

4.1.3 Random coefficient models

As a means of providing an adequate measure of the underlying relationship between two accounting variables, models estimated by OLS provide useful modelling frameworks. However, it has been demonstrated elsewhere (with an indirect test based on published summary statistics) that the sign and magnitude of the constant term will vary not only from one financial ratio to another, but from one sector to another and from one size of company to another (McLeay and Fieldsend, 1987). Hence, the model introduced in Equation 4.7 is often referred to as the fixed effects model as it does not allow cross-sectional variations in the parameters to be considered (Fig. 4.2).

The fixed effects model assumes that random effects are associated only with companies. However, it seems reasonable to conjecture that there is greater homogeneity of companies' ratios within the industrial sector. The classical way of dealing with this is to extend Equation 4.7 to an analysis of covariance model, with the sector as a classifying factor. In cross-sectional performance studies, however, such an approach is considered unsatisfactory because of the large number of sectors, and hence parameters (Fieldsend *et al.*, 1987). Furthermore, industry data are often unbalanced, with some sectors containing only a small number of companies. It is for these reasons, and

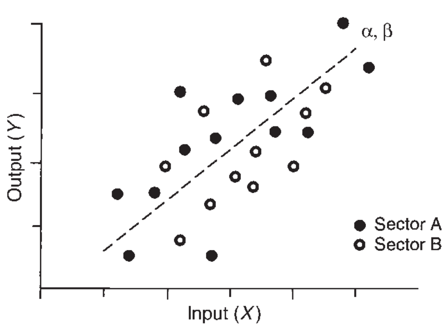


Fig. 4.2. Fixed effect regression model fails with cross-sectional data.

particularly in order to achieve parsimony in model building, that variance component analysis has been used.

Variance components

Marketing researchers frequently encounter cross-sectional data when developing sales response models. One appropriate approach to analysing such data is to estimate a separate OLS equation for each cross-section. Alternatively, one could pool the data from all cross-sections to estimate a single set of response coefficients for all cross-sections (Leone *et al.*, 1993). However, when data are pooled, the responsiveness of individual cross-sections cannot be evaluated.

Variance component models can be treated as generalizations of ordinary regression models (Harville, 1977). Among a number of applications, they are appropriate for unbalanced data having a hierarchical structure where total variability may be separated into components attributable to each 'cluster' of observations (Goldstein, 1986). In this case, Fieldsend *et al.* (1987) are suggesting that companies cluster into industrial sectors and, although there might be great variability in a particular financial indicator over the broad cross-section of companies, it is reasonable to expect some similarity in financial characteristics of companies within sectors (Lee, 1985). A variance component model will take into account this 'within-sector' homogeneity.

In this respect, variance component analysis treats the variation of sector effects as a component of the unexplained variability, the overall variance being partitioned into components for 'sector' as well as the 'company residual'. Introducing a term associated with the sector in Equation 4.7 gives

$$\ln Y_{ij} = \alpha + \beta \ln X_{ij} + \gamma_j + \varepsilon_i \tag{4.8}$$

which is called the mixed effects model, where the indices *i* and *j* represent company *i* in sector *j*. The regression coefficient α now represents the intercept for the 'average' sector, as the unexplained variability is now represented by two random terms – the sector effects γ_j and the company effects ε_i, where γ and ε are mutually independent and normally distributed with the means 0 and variances $\sigma^2{}_{\mathrm{I}}$ and σ^2. Thus, for the *i*th company in the *j*th sector, the model for $\ln Y$ is composed of the following terms.

1. An industry-specific effect $\alpha_j = \alpha + \gamma_j$.
2. $\beta \ln X_{ij}$, where β is constant over all sectors.
3. A residual company effect, ε_i.

In other words, while the slope in this mixed effects model remains constant with β as a fixed effect, the intercept α_j varies from sector to sector as a random effect (Fig. 4.3).

Of course, the slope of $\ln X$ may also be allowed to vary, giving the model

$$\ln Y_{ij} = \alpha + \beta \ln X_{ij} + \gamma_j + \delta_j \ln X_{ij} + \varepsilon_i \tag{4.9}$$

where δ_j is assumed to be a random normal variate with zero mean and variance $\sigma^2{}_s$. In the literature, Equation 4.9 is referred to as the random effects model. An alternative presentation is given by

$$\ln Y_{ij} = \alpha_j + \beta_j \ln X_{ij} + \varepsilon_i \tag{4.10}$$

where the random slope $\beta_j = \beta + \delta_j$ and the random intercept $\alpha_j = \alpha + \gamma_j$ each vary from sector to sector. The interpretation of this model is, firstly, that a high α_j will indicate a relatively high value of $\ln Y$ proportionate to $\ln X$ in sector j and, secondly, that a large difference between β_j and β will indicate a relatively greater size effect in that sector.

The fixed effects model (Equation 4.7), the mixed effects model (Equation 4.8) and the random effects model (Equation 4.10) are illustrated in Figs 4.2, 4.3 and 4.4, respectively. In fitting the fixed effects model, no account is taken of the differences in the two sectors at this stage, giving therefore the single regression line (α, β) that is superimposed on the plot of data points.

In Figs 4.3 and 4.4, the fitted lines that are obtained with Equations 4.8 and 4.10 can be compared. Parameter estimates from the 'between-sector' mixed effects model provide the regression lines in Fig. 4.3, where the slope remains constant but the intercept varies between the two sectors. Here, there are two fitted lines α_A, β and α_B, β for sectors A and B, respectively, showing clearly the differences between the two groups in the intercept. Situated between these fitted lines is the regression line for the 'average' sector (α, β) where, as noted before, the estimates of α and β differ from the OLS estimates

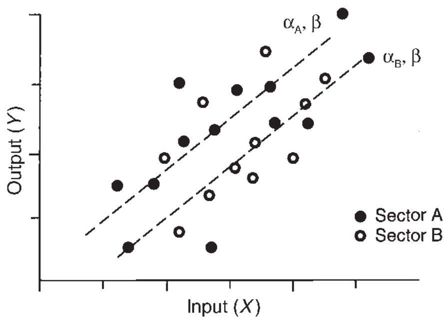


Fig. 4.3. The mixed effects regression model.

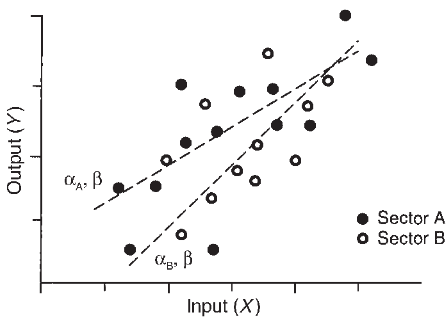


Fig. 4.4. The random effects regression model.

obtained with the fixed effects model (Equation 4.7). But, looking at the data suggests that there is likely to be some difference in the slopes, reflecting the within-sector variation in the size effect. Fitting the 'within-sector' random effects model (Equation 4.10), gives the plot in Fig. 4.4. Here the convergence of the fitted lines for sectors A and B (α_A, β_A and α_B, β_B) is evident.

Leone *et al.* (1993) introduce a version of the random coefficient model that can be used to estimate separate sets of response coefficients for each cross-section, thereby circumventing the assumption that coefficients are homogeneous in all cross-sections. They demonstrate this approach with an empirical model that relates brand level sales to price and advertising.

4.1.4 Mixture models

In marketing, the fact that it is not very difficult to find reasons for the potential presence of heterogeneous groups or segments has led to the wide use of clustering and, more recently, of unconditional mixture procedures for market segmentation. In those applications, the main objective is descriptive, that is, to form homogeneous groups of consumers on the basis of several observed characteristics. More recently, attention has shifted towards forming segments that are homogeneous in terms of their responsiveness to price, sales promotions, etc. Such segments are formed on the basis of the inferred relationship between a response variable measuring behaviour (purchase volume, brand choice, etc.) intentions, or stated preferences and a set of causal variables (product features, price, sales promotion, etc.) within each homogeneous group.

Mixture models enable marketing researchers to cope with heterogeneity in their samples and to identify segments by using a model-based approach that connects classical clustering to conventional statistical estimation methods. One of the first examples of mixed analysis of variance models was given by Hartley and Rao (1967), and was later expanded by Everitt and Hand (1981) and Titterington *et al.* (1985). More recently, Wedel and Kamakura (1999) provided a general framework for market segmentation based on mixture models.

In the basic unconditional mixture models, also called finite mixture models, it is assumed that a sample is composed of a number of underlying groups or segments. 'Unconditional' refers to the situation in which there are no exogenous variables explaining the means and variances of each component in the unconditional mixture distribution. The unconditional mixture models correspond to descriptive clustering approaches, e.g. the K-means procedure. In the unconditional mixture of normal distributions, for example, the mean and variance of each underlying segment is estimated directly.

In the unconditional mixture model, in order to describe the process generating the measurements, a certain statistical distribution is assumed for each of the groups. Such a distribution function describes the probabilities that the measurements take certain values. The most commonly used distribution function is the normal distribution, other distribution functions include the binomial, the Poisson and the exponential distribution.

Given one of these assumed distributional forms, the purpose of the mixture approach is to decompose the sample into underlying groups. Unconditional mixture models have received wide application, being used to model distributions where the measurements arise from separate groups, but individual membership is unknown. As methods of density estimation, mixture models provide more flexible models than the simple normal-based models (e.g. linear discriminant analysis) providing improved discrimination in some circumstances.

There are several issues associated with mixture models that are of interest. The most important for density estimation concerns the estimation of the model parameters. Whereas in the classical approaches for unconditional mixtures only the expected values of each of the underlying densities are estimated, Wedel and DeSarbo (1995) proposed a mixture methodology that enables the estimation of the relation of the observations in each underlying group, with a set of explanatory variables. They implemented conditional mixture models, also called mixture regression models, in an easy to use computer program. Conditional mixture models allow for a probabilistic classification of observations into segments and simultaneous estimation of a generalized linear regression model within each segment.[2]

A complete formulation of the unconditional and conditional mixture model can be found in Wedel and Kamakura (1999). In the next two sections an illustrative example for the unconditional mixture model and an application of the EM algorithm in an Excel spreadsheet are given. Finally, the differences between the unconditional and conditional mixture model will be highlighted.

Unconditional mixture models

The unconditional mixture model approach assumes that objects (from now on referred to as 'firms') on which the variables y_{nk} are measured, arise from a population that is a mixture of S groups (from now on referred to as 'segment' (cross-section) of an industry), in proportions $\pi_1, \ldots, \pi_S$. In the beginning it is not known to which segment a particular firm belongs. The probabilities π_s are subject to the following constraints.

$$\sum_{s=1}^{S} \pi_s = 1,\ \pi_s \geq 0,\ i = 1, \ldots, I \tag{4.11}$$

[2] The mixture approach to clustering provides a flexible class of clustering algorithms that can be tailored to a very wide range of substantive problems. Although Wedel and Kamakura (1997) and Wedel and DeSarbo (1995) apply mixture models in the field of consumer research, their applications to consumer panel data can easily be adopted by business performance research applications with company panel data sets.

Given that y_{nk} comes from segment s, the conditional distribution function of the vector y_n is represented by the general form $f_s(y_n | \theta_s)$, where θ_s denotes the vector of all unknown parameters associated with the specific form of the density chosen. For example, in the case that y_{nk} within each segment are independently normally distributed, θ_s contains the means, μ_{ks}, and variances, σ_s^2, of the normal distribution within each of the S segments. The simple idea behind the mixture distributions is that if the distributions conditional upon knowing the segments have been formulated, the unconditional distribution of y_n is obtained as

$$f(y_n | \phi) = \sum_{s=1}^{S} \pi_s f_s(y_n | \theta_s) \tag{4.12}$$

where $\phi = (\pi, \theta)$. Thus, from probability theory it can be derived that the unconditional probability is equal to the product of the conditional probability, given s times the probability of s, and that expression summed over all values of s.

In the unconditional mixture method, each different group S in the population is assumed to be described by a different probability distribution. These different probability distributions may belong to the same family but differ in the values they take for the parameters of the distributions. Many forms of mixture distributions can be considered and there are many methods for estimating their parameters. Wedel and Kamakura (1999) developed mixture models for the most commonly used distributions, which are members of the exponential family of distributions.[3]

The purpose of the analysis of mixture models is to estimate the parameter vector $\phi = (\pi, \theta)$. This is accomplished by using the method of maximum likelihood (ML), which provides a statistical framework for assessing the information available in the data about the parameters in the model. The likelihood for ϕ is formulated as:

$$L(\phi; y) = \prod_{n=1}^{N} f(y_n | \theta_s) \tag{4.13}$$

An estimate of ϕ can be obtained by maximizing the above likelihood equation (4.13) with respect to ϕ being subject to the restrictions in Equation 4.11. The purpose of ML estimation is to find a parameter vector ϕ_0 such that the observations y are more likely to have come from $f(y | \phi_0)$ than from $f(y | \phi)$ for any other value of ϕ. The estimate ϕ_0 thus maximizes the likelihood of the observed data, given the model. This is accomplished by maximizing the likelihood function $L(\cdot)$ above, which measures the likelihood that the parameter vector ϕ could have produced the observed vector y. The likelihood is simply the product over the densities of the N individual observations (or observation vectors), as those are assumed to be independent.

3 The well-known normal, Poisson and binomial distributions, as well as the negative binomial, Dirichlet, exponential gamma and inverse Gaussian distributions, belong to this group and their properties are explained in detail by Wedel and Kamakura (1997: 78).

In general, it is not possible to solve $L(\cdot)$ explicitly for the parameters of the model, so iterative schemes are employed. In this context, among other numerical methods, the Fisher scoring algorithm and the EM algorithm are the most popular algorithms for estimation. The Fisher scoring algorithm has been implemented in a computer program by Longford (1986, 1987). A discussion of the algorithm's characteristics and applicability appears in Aitkin and Longford (1986). An alternative approach for maximizing the likelihood $L(\cdot)$ is to use a general class of iterative procedures known as EM (expectation–maximization) algorithms, introduced in the context of missing data estimation by Dempster *et al.* (1977) and explained below.

THE EM ALGORITHM The EM algorithm has apparently been the most popular ML estimation algorithm because of its computational simplicity. The basic idea behind the EM algorithm is the following: rather than apply a complicated maximization routine to maximize the likelihood across the entire parameter space, one augments the observed data with additional information that replaces unobserved data, greatly simplifying the maximization of the likelihood. In the context of mixture models, these unobserved data constitute the membership of subjects in segments. The membership is unobserved, but were it observed, it would facilitate the computation of the segment-level estimates. The information that is substituted for the missing data is the expected membership of subjects, given a set of preliminary estimates of the parameters of the model.

To derive the EM algorithm for mixture models, z_{ns} is introduced as unobserved data, indicating whether observation n belongs to mixture component s: $z_{ns} = 1$ if n comes from segment s, and $z_{ns} = 0$ otherwise. The z_{ns} values are assumed to be independently multinomially distributed, arising from one draw of the S segments with probabilities π_s. Further, the observed data y_n are assumed to be independent. With z_{ns} considered as missing data, the log-likelihood function for the complete data X and $Z = (z_{ns})$ can be formed:

$$\ln L_c(\phi) = \sum_{n=1}^{N} \sum_{s=1}^{S} \left\{ z_{ns} \ln f_s \left(y_n \,|\, \theta_s \right) + z_{ns} \ln \pi_s \right\} \tag{4.14}$$

That complete log-likelihood is maximized by using an iterative EM algorithm.

In the E-step of the EM algorithm, the expectation of log-likelihood is calculated with respect to the unobserved data Z, given the observed data Y and the provisional estimates of ϕ. This expectation is obtained by replacing Z in the likelihood by their expected values, which is identical to the posterior probability that object n belongs to segment s (p_{ns}). The posterior probability derived by means of Bayes' Theorem is given by[4]

[4] The posterior probabilities p_{ns} play a crucial role in the estimation of the parameters for each segment, so that the more information that is available from each firm n, the more accurate the segment-level estimates will be.

$$p_{ns} = \frac{\pi_s f_s(y_n|\theta_s)}{\sum_{s=1}^{S} \pi_s f_s(y_n|\theta_s)} \tag{4.15}$$

In the M-step of the EM algorithm, the estimates of the prior probabilities at each step of the algorithm are simply the averages of the posterior probabilities in each segment, which is:

$$\hat{\pi}_s = \frac{1}{N} \sum_{n=1}^{N} \frac{\hat{p}_{ns}}{N} \tag{4.16}$$

The new maximum likelihood equations θ_s are obtained by weighting the contribution from each sample unit with the posterior probabilities of segment membership.

An attractive feature of the EM algorithm is that the above equations have a closed form for some models. For example, if $f_s(y_{nk}|\theta_s)$ is the density of the normal or Poisson distributions, the mixture component means are estimated by

$$\mu_s = \frac{\sum_{n=1}^{N} p_{ns} y_n}{\sum_{n=1}^{N} p_{ns}} \tag{4.17}$$

which is the weighted mean of the observations.[5] The E- and M-steps are alternated until the improvement in the likelihood function gets marginal and the user decides to stop the procedure.

An attractive feature of the EM algorithm is that it provides monotonically increasing values of the log-likelihood. Figure 4.5 illustrates the EM iterative

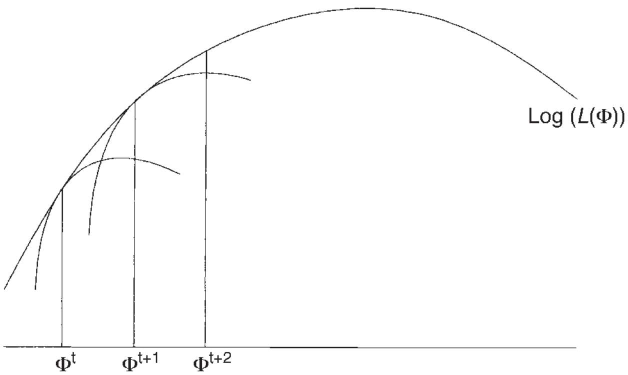


Fig. 4.5. Illustration of the EM algorithm.

[5] Similarly, the component variances for the normal distribution are calculated by using the posterior membership probabilities as weights.

scheme. There one sees that a local maximum of a function $f(\phi)$ is found by defining an auxiliary function $f'(\phi^t, \phi^{t+1})$ that touches the function f at the point (ϕ^t, ϕ^{t+1}) and lies everywhere else below it. The auxiliary function is maximized and the position of the maximum, ϕ^{t+1}, gives a new value of the original function f that is greater than at the previous iteration. This process is repeated, with a new auxiliary function being defined that touches the curve at $(\phi^{t+1}, f(\phi^{t+1}))$ and continues until convergence. The shape as well as the position of the auxiliary function will also change as the iteration proceeds (see Dempster *et al.*, 1977 for details). Under moderate conditions, the log-likelihood is bound from above, thus assuring that the algorithm converges to at least a local optimum.

The EM algorithm provides a general framework for estimating a large variety of mixture models. The M-step amounts to the maximum likelihood estimation of the segment-level model for each segment s, in which the contribution by each firm n to the likelihood is weighted by the posterior probabilities p_{ns} obtained in the E-step. The EM algorithm makes the extension of aggregate models to the segment level quite straightforward, as long as maximum-likelihood estimates can be obtained for the aggregate model either analytically or through numerical optimization. The estimation for those models involves only modifications of the likelihood equations in the M-step, weighting each firm by the posterior probabilities computed in the E-step.

AN UNCONDITIONAL MIXTURE MODEL EXAMPLE The EM algorithm for unconditional mixture models is a specific instance of the EM algorithm for conditional mixture regression models. Figure 4.6 provides a simple example of the EM algorithm implemented in an Excel spreadsheet. The data generated for 20 hypothetical subjects come from a two-segment Poisson model. In the context of business performance research, this example could refer to 20 firms from two cross-sections of an industry producing, for instance, a financial key ratio with distinct levels for each of the two segments. For this demonstration the firms 1 to 10 are in segment 1 with a key ratio of $\mu_1 = 1$ and firms 11 to 20 are in segment 2 with a key ratio of $\mu_2 = 10$. The data generated are displayed in column B of Fig. 4.6. The EM algorithm is applied to this hypothetical small data set to illustrate the procedure. The procedure assumes $s = 2$, which should lead the EM algorithm to the true number of segments. Figure 4.6 displays the two initial steps of the algorithm, which starts with the random generation of the first estimated probabilities listed in columns C and D, respectively.

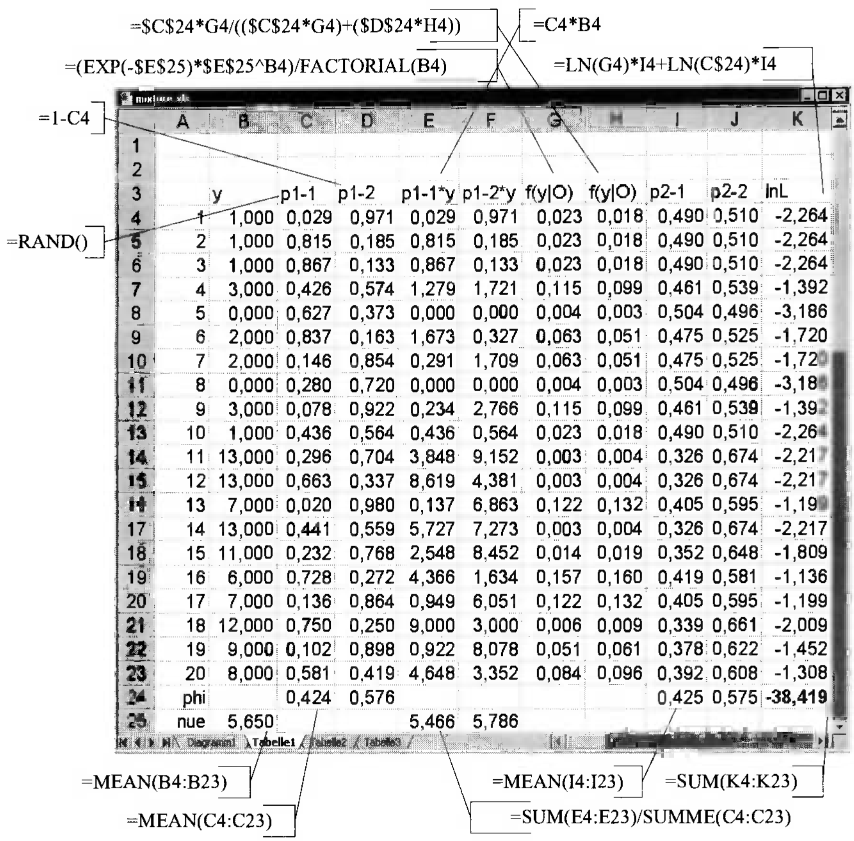
Given the initial posterior values, the EM algorithm computes the segment means in the M-step (see cells C24 and D24). The M-step for this model entails only closed-form expressions. The estimates of the prior probabilities π are equal to the mean of the posteriors for each segment, according to Equation 4.16, and the segment means are, according to Equation 4.17, simple weighted means of the data, where the weights constitute the posterior probabilities at step 1 (see cells E25 and F25).

In step 2, new posteriors (see columns I and J) are calculated from the current estimates of the priors (see columns G and H) and the segment means, according to Equation 4.15 and the Poisson distribution, which is

$$\frac{e^{-\mu}\mu^{y}}{y!} \tag{4.18}$$

In the case of the Poisson distribution, the kernel of log-likelihood takes a simple form and equals $y_n \ln(\mu_s) - \mu_s$. Thus, the posteriors in step 2 are calculated on the basis of the segment means and the posteriors in columns I and J. New segment means are estimated as a weighted average on the basis of the new priors in step 2, and so on.

In this example the structure of the data is already well recovered after a few iterations because the segments are very well separated. Table 4.1 provides a summary of the EM algorithm, which converges after eight iterations, showing that the change in the likelihood from step 7 to step 8 was less



	A	B	C	D	E	F	G	H	I	J	K
1											
2											
3		y	p1-1	p1-2	p1-1*y	p1-2*y	f(y|O)	f(y|O)	p2-1	p2-2	lnL
4	1	1,000	0,029	0,971	0,029	0,971	0,023	0,018	0,490	0,510	-2,264
5	2	1,000	0,815	0,185	0,815	0,185	0,023	0,018	0,490	0,510	-2,264
6	3	1,000	0,867	0,133	0,867	0,133	0,023	0,018	0,490	0,510	-2,264
7	4	3,000	0,426	0,574	1,279	1,721	0,115	0,099	0,461	0,539	-1,392
8	5	0,000	0,627	0,373	0,000	0,000	0,004	0,003	0,504	0,496	-3,186
9	6	2,000	0,837	0,163	1,673	0,327	0,063	0,051	0,475	0,525	-1,720
10	7	2,000	0,146	0,854	0,291	1,709	0,063	0,051	0,475	0,525	-1,72[illegible]
11	8	0,000	0,280	0,720	0,000	0,000	0,004	0,003	0,504	0,496	-3,18[illegible]
12	9	3,000	0,078	0,922	0,234	2,766	0,115	0,099	0,461	0,539	-1,39[illegible]
13	10	1,000	0,436	0,564	0,436	0,564	0,023	0,018	0,490	0,510	-2,26[illegible]
14	11	13,000	0,296	0,704	3,848	9,152	0,003	0,004	0,326	0,674	-2,21[illegible]
15	12	13,000	0,663	0,337	8,619	4,381	0,003	0,004	0,326	0,674	-2,21[illegible]
16	13	7,000	0,020	0,980	0,137	6,863	0,122	0,132	0,405	0,595	-1,19[illegible]
17	14	13,000	0,441	0,559	5,727	7,273	0,003	0,004	0,326	0,674	-2,217
18	15	11,000	0,232	0,768	2,548	8,452	0,014	0,019	0,352	0,648	-1,809
19	16	6,000	0,728	0,272	4,366	1,634	0,157	0,160	0,419	0,581	-1,136
20	17	7,000	0,136	0,864	0,949	6,051	0,122	0,132	0,405	0,595	-1,199
21	18	12,000	0,750	0,250	9,000	3,000	0,006	0,009	0,339	0,661	-2,009
22	19	9,000	0,102	0,898	0,922	8,078	0,051	0,061	0,378	0,622	-1,452
23	20	8,000	0,581	0,419	4,648	3,352	0,084	0,096	0,392	0,608	-1,308
24	phi		0,424	0,576					0,425	0,575	**-38,419**
25	nue	5,650			5,466	5,786					

Fig. 4.6. First two steps of an unconditional mixture model example.

than 0.001%. Columns two and three of Table 4.1 provide the value of the log-likelihood during the iterations. Observe that it decreases considerably during the first three steps, but the subsequent decrease is much smaller. The value of the log-likelihood was −21.164 at convergence. At the final step of the iterations, the true means appear to be quite well recovered despite the small number of observations.

Conditional mixture models

The identification of segments and simultaneous estimation of the response functions within each segment has been accomplished by a variety of conditional mixture models, including mixtures of linear regressions (DeSarbo and Cron, 1988), multinomial logits (Kamakura and Russell, 1989), rank logits (Kamakura *et al.*, 1994) and generalized clusterwise regression models (Lau *et al.*, 1999). Most importantly, these conditional mixture models directly identify segments to actionable normative segmentation.

Whereas the major thrust of development and application of conditional mixture models has been in marketing and business research, there is potential for substantive applications in virtually all physical and social sciences.

Table 4.1. Summary of EM-optimization after eight iterations.

i	LnL	LnL^t – LnL^{t-1}	n	p_{81}	p_{82}
1	–38.418853		1	0.99837735	0.00162265
2	–37.087361	–1.331492	2	0.99837735	0.00162265
3	–30.318358	–6.769003	3	0.99837735	0.00162265
4	–21.826537	–8.491821	4	0.92614464	0.07385536
5	–21.172308	–0.654229	5	0.99976802	0.00023198
6	–21.161378	–0.010930	6	0.98874357	0.01125643
7	–21.163315	0.001937	7	0.98874357	0.01125643
8	–21.163865	0.000550	8	0.99976802	0.00023198
			9	0.92614464	0.07385536
			10	0.99837735	0.00162265
			11	0.00000004	0.99999996
			12	0.00000004	0.99999996
			13	0.00518199	0.99481801
			14	0.00000004	0.99999996
			15	0.00000216	0.99999784
			16	0.03520261	0.96479739
			17	0.00518199	0.99481801
			18	0.00000031	0.99999969
			19	0.00010615	0.99989385
			20	0.00074309	0.99925691
			π_s	0.49346201	0.50653799
			μ_s	1.39803058	9.79270540

Applications of generalized linear models, which include, as special cases, linear regression, logit and probit models, log-linear and multinomial models, inverse polynomial models, and some models used for survival data, have been numerous.

In analogy to the previous section on unconditional mixture models, the following conditional mixture model within the exponential family is formulated (Wedel and DeSarbo, 1995). Assume the vector of observations on object n, y_n, arises from a population that is a mixture of s segments in proportions $\pi_1, \ldots, \pi_s$, where the segment from which a particular vector of observations arises is unknown. The probabilities π_s are positive and sum to one as in Equation 4.11. Assume that the distribution of y_n, given that y_n comes from segment s, $f_s(y_{nk} | \theta_s)$, is one of the distributions in the exponential family or the multivariate exponential family. Conditional on segment s, the y_n are independent. If y_{nk} cannot be assumed independent across k (i.e. repeated measurement on each subject or firm), a distribution within the multivariate exponential family, such as the multivariate normal or the multinomial, is appropriate. The distribution $f_s(y_{nk} | \theta_s)$ is characterized by parameters θ_{sk}. The means of the distribution in segment s (or expectations) are denoted by μ_{sk}. Some of the distributions, such as the normal, also have an associated dispersion parameter λ_s that characterizes the variance of the observations within each segment.

The development of the model is very similar to that of the mixture models described earlier. A major difference, however, is that the means of the observations in each segment are being predicted by using a set of explanatory variables, as in the unconditional mixture models. Next, a linear predictor η_{nsk}, is specified which is produced by P explanatory variables $X_1, \ldots, X_P$ ($X_p = (X_{nkp})$; $p = 1, \ldots, P$) and parameter vectors $\beta_s = (\beta_{sp})$ in segment s:

$$\eta_{nsk} = \sum_{p=1}^{P} X_{nkp} \beta_{sp} \tag{4.19}$$

The linear predictor is thus a linear combination of the explanatory variables and a set of coefficients that are to be estimated. The linear predictor is in turn related to the mean of the distribution, μ_{sk}, through a link function $g(\cdot)$ such that in segment s:

$$\eta_{nsk} = g(\mu_{nsk}) \tag{4.20}$$

Thus, for each segment, a generalized linear model is formulated with a specification of the distribution of the variable (within the exponential family), a linear predictor η_{nsk} and a function $g(\cdot)$ that links the linear predictor to the expectation of the distribution. For each distribution there are preferred links, called canonical links. The canonical links for the normal, Poisson, binomial, gamma and inverse Gaussian distributions are the identity, log, logit, inverse and squared inverse functions, respectively. For example, for the normal distribution the identity link involves $\eta_{nsk} = \mu_{sk}$, so that by combining Equations 4.19 and 4.20, the standard linear regression model with segments arises.

The unconditional probability density function of an observation vector y_n can now be expressed in the unconditional mixture form:

$$f(y_n|\phi) = \sum_{s=1}^{S} \pi_s f_s(y_n|\theta_s) \tag{4.21}$$

where the parameter vector $\phi = (\pi_s, \theta_s)$ and $\theta_s = (\beta_s, \lambda_s)$. Note that the difference from Equation 4.13 is that here the mean vector of the observations (given each segment) is parameterized in terms of regression coefficients relating the means to a set of explanatory variables.

In the EM estimation the purpose is to estimate the parameter vector ϕ. Again, the problem can be solved using the EM algorithm as described in the preceding section.

The EM algorithm's application is very similar to that described for ordinary mixture models, except for the estimation of the within-class generalized linear model. Once an estimate of ϕ has been obtained, the posterior membership probabilities p_{ns} are calculated in the E-step, as shown in Equation 4.15. In the M-step, the expectation of $\ln L_c$ is maximized with respect to ϕ to obtain new provisional estimates. For the complete formula of the EM algorithm in the conditional case of a generalized mixture model refer to Wedel and Kamakura (1999). An application of the conditional mixture model for the selection of comparison partners in business performance studies is given in Chapter 6.

Problems with mixture models

The EM algorithm is computationally attractive as it can be programmed easily and convergence to a local optimum is ensured. However, there are several severe problems in conjunction with mixture models which have to be addressed when the technique is used for the search of comparison partners in business performance studies. These problems arise from limitations associated with the EM algorithm, model identification and the specification of the appropriate number of segments.

LIMITATIONS OF THE EM ALGORITHM In comparison with direct numerical optimization of the likelihood, the convergence of the EM algorithm can be slow, depending on the data distribution and the initial estimates for the parameters (Titterington *et al.*, 1985). The computation time needed to reach convergence increases quadratically by the number of parameters, thus sometimes more than 100 iterations are necessary only for a few parameters.

Another potential problem associated with the application of the EM algorithm to mixture problems is its convergence to local optima. The estimation procedure is sensitive to local optima, and the algorithm may end up in one of those optima depending on the values of the parameters used to start the iterative search process (Titterington *et al.*, 1985: 84). Convergence to local optima occurs in particular: (i) if segments are not very well separated; (ii) if the number of parameters estimated is large; and/or (iii) if the information embedded in each observation is limited, leading to a relatively weak posterior update of the membership probabilities (p_{ns}) in the E-step. The likelihood of accepting a local optimum can be reduced by starting the EM algorithm from a wide range of (random) starting values or from a larger number of classes. Another way to avoid accepting a local optima convergence is to use some clustering procedure such as K-means, and apply it to the dependent variable to obtain an initial partition of the data. In any case, it is worthwhile to try several initializations, since agreement between the resulting classifications lends more weight to the chosen solution. If different starting values yield different optima, the solution with the maximum value of the likelihood is recommended as the best solution (Wedel, 1997).

When to stop the algorithm is also an issue. The EM algorithm is stopped when the likelihood changes less than some small amount from one iteration to the next. However, some researchers have argued that this is a measure of lack of progress rather than a measure of convergence and there is evidence that the algorithm is often stopped too early. The issue here is that the point of stopping may be well off the optimum values.

PROBLEM OF IDENTIFICATION Identification of a model refers to the situation in which only one set of parameter values uniquely maximizes the likelihood. The model is said to be unidentified when more than one set of values result in a maximum. The interpretation of parameters from unidentified models is useless because an infinite set of parameters yields the same solution.

Problems of identification exist in conditional mixture models. Conditional mixture models have a particular identification problem: identification related to the condition of the predictor matrix X. As in linear models, collinearity among predictors may lead to problems of identification, resulting in unstable estimates of the regression coefficients and large standard errors. In conditional mixture models that situation is compounded by the fact that there are fewer observations for estimating the regression model within each segment than there are at the aggregate level. Therefore identification related to the condition of the X matrix is an important issue in applications of conditional mixture models.

IDENTIFYING THE APPROPRIATE NUMBER OF SEGMENTS A major difficulty with the method of mixture models relates to the number of segments. Many algorithms require the number of segments to be specified before the remaining parameters can be estimated. Several test statistics have been put forward. However, when applying the above models, the actual number of segments is unknown and must be inferred from the data. Unfortunately, a way of identifying the number of segments is still without satisfactory statistical solution. For example, suppose someone wants to test the null hypothesis (H_0) of s segments against the alternative hypothesis (H_1) of $s + 1$ segments. The likelihood ratio statistic[6] used in many ML estimations does not apply, since the statistic is not asymptotically distributed as chi-square. A procedure for determining the number of segments is the Monte Carlo test procedure. Because of the high computational burden of the Monte Carlo procedure, another approach for investigating the number of segments is frequently used. In this case there is an attempt to balance the increase in fit obtained against the number of parameters estimated. For models with more segments there are more parameters. Basically, a penalty is imposed on the likelihood that is related to the number of parameters estimated:

$$C = 2\ln L + Pd \tag{4.22}$$

P is the number of parameters estimated and d is a constant. The constant imposes a penalty on the likelihood, which weights the increase in fit (more parameters yield a higher likelihood) against the additional number of parameters estimated. The classical Akaike (1973) information criterion, AIC, arises when $d = 2$. Two criteria that penalize the likelihood more heavily are the Bayesian information criterion, BIC, and the consistent Akaike information criterion, CAIC. For those criteria, $d = \ln(N)$ and $d = \ln(N + 1)$, respectively. Note that the CAIC penalizes the likelihood even more than BIC, although the two criteria are quite close. Both statistics impose an additional sample size penalty on the likelihood, and are more conservative than the AIC statistic in that they tend to favour models with fewer segments. Studies by Bozdogan (1994) indicate that CAIC is preferable in general for mixture models. Bozdogan also proposed the modified AIC (MAIC), for which $d = 3$. The major problem with these criteria is that they rely on the same properties as the likelihood ratio test, and can therefore be used only as indicative of the number of segments.

Another approach, the non-parametric maximum likelihood estimation (NPMLE), treats the number of segments in conditional mixture models as a random variable rather than a fixed number (Dillon and Kumar, 1994).

4.2 Frontier Models

Since the publication of Aigner *et al.* (1977), the field of frontier estimation and efficiency measurement has seen a rapid and parallel development in the econometric and operations research/management science literatures. The statistical estimation of the production function was followed by the econometric research community, whereas the operations research/management science approach developed out of mathematical programming and goes by

[6] The likelihood ratio test statistic is simply the difference of the maximum likelihood of the two models. The statistic follows a chi-square distribution for nested models under the null hypothesis, given that certain regularity conditions are satisfied.

the descriptive term of data envelopment analysis (DEA). For many years these two approaches have proceeded independently of one another.

The two approaches share a focus on extreme observations implied, and extracting information from them. They have somewhat different practical motivations. DEA is geared toward the managerial implications of efficiency measurement, particularly in the public sector where output prices often cannot be specified. DEA uses different techniques to construct frontiers, and deviations from constructed frontiers are interpreted very differently.

Work on DEA was not universally accepted by economists who, in general, continued to develop parametric methods. In 1986, Grosskopf describes how these developments 'have led many economists to believe that the non-parametric approach is obsolete, largely because of the restrictions placed on the technology in the early studies employing that approach' (Grosskopf, 1986). Grosskopf proceeds to point out that the non-parametric approach is much more flexible than had been suggested and that it has been underestimated by economists. Notwithstanding this debate between economists and management scientists, some progress towards the convergence of the two approaches and some indication of their complementarity have been observed recently. Selected papers presented during the first joint conference on 'Parametric and Nonparametric Approaches to Frontier Analyses' held at the University of North Carolina at Chapel Hill in the autumn of 1988 are summarized in a special issue of the *Journal of Econometrics*, edited by Lewin and Lovell (1990), and reflect the differences that exist. In the theory/modelling arena, ongoing work in pursuit of a convergence of the two techniques can be expected to expand, both by making DEA stochastic and by relaxing parametric restrictions in econometric models.

4.2.1 Data envelopment analysis

DEA is a non-parametric technique, i.e. it can be used to compare input/output data making no prior assumptions about the shapes of the probability distributions under study. The breakthrough of DEA came in the research work undertaken by Charnes *et al*. (1978). If Farrell's 1957 paper is taken as the seminal work, the research reported in 1978 is undoubtedly the basis for all subsequent developments in the non-parametric approach to evaluating technical efficiency. Until today approximately 2000 research articles and reports have been published on DEA which can be found predominantly in Operations Research literature.[7] Recently, two special issues by *Annals of Operations Research* (vols 66 and 73) and one by *European Journal of Operational Research*

[7] For example *European Journal of Operational Research, Journal of Operational Research, Journal of the Operational Research Society, Annals of Operations Research, Management Science, Omega – International Journal of Management Science, Research in Governmental and Nonprofit Accounting.*

(vol. 115) provide a review of the evolution, development and future research directions on the use of DEA (Cooper *et al.*, 1996; Lewin and Seiford, 1997). During the last 20 years several textbooks have appeared. The most widely used are Norman and Stoker (1991), Ganley and Cubbin (1992), Charnes *et al.* (1994b), Sengupta (1995), Avkiran (1999) and Cooper *et al.* (2000).

DEA has been used to assess productivity in a number of types of applications. The focus of the original work of Charnes *et al.* (1978) was on decision making by ‘not-for-profit’ entities. It concentrated on multifactorial problems (particularly with reference to outputs) and could discount economic weighting factors such as market prices. The majority of empirical studies are in the fields of medical services (Nyman and Bricker, 1989; Morey *et al.*, 1990; Valdamis, 1992; Banker *et al.*, 1998; Löthgren and Tambour, 1999), educational institutions (Charnes *et al.*, 1981; Tomkins and Green, 1988; Ahn and Seiford, 1990; Ray, 1991; Doyle *et al.*, 1996; Hanke and Leopoldseder, 1998; Sarrico and Dyson, 1998; Grosskopf *et al.*, 1999; Bifulco and Bretschneider, 2001) and in other forms of public authorities or services (Lewin and Morey, 1981; Charnes *et al.*, 1985a; Nunamaker, 1985; Banker, 1989; Schinnar, 1990; Ganley and Cubbin, 1992; Thanassoulis, 1995; Ruggiero, 1996; Worthington, 2000). However, since 1978 numerous researchers have shown that the DEA approach is applicable to the private as well as the public sector. Various applications can also be found for the profit-oriented industries; however, these are more frequently for the purpose of internal than external benchmarking. Relatively rarely publications on DEA applications appear for manufacturing industries (Kamakura *et al.*, 1988; Schefczyk, 1993; Schefczyk and Gerpott, 1995; Hawdon and Hodson, 1996; Westermann, 1996; Chandra *et al.*, 1998; Al-Shammari, 1999; Caporaletti *et al.*, 1999); more frequently applications can be found for industries in various service sectors, e.g. for the measuring of efficiency of banks (Charnes *et al.*, 1990; Ferrier and Lovell, 1990; Barr *et al.*, 1993, 1994; Yeh, 1996; Barr and Siems, 1997; Siems and Barr, 1998; Golany and Storbeck, 1999; Kantor and Maital, 1999; Maital and Vaninsky, 1999; Soteriou and Zenios, 1999; Sueyoshi, 1999; Thanassoulis, 1999; Zenios *et al.*, 1999), airports (Adler and Berechman, 2001; Martin and Román, 2001; Pels *et al.*, 2001), retail stores (Thomas *et al.*, 1998), mutual funds (Morey and Morey, 1999), investments in information technology (Shafer and Byrd, 2000), benchmarking of computer hard- and software (Doyle and Green, 1994; Jammernegg *et al.*, 1997) and the control of electricity power plants (Athanassopoulos *et al.*, 1999).

The variety of applications clearly shows the wide appeal of DEA. However, although efficient frontier methods have been used extensively in research found in other literature, there has been little research that examines efficiency in the tourism industry using DEA. In 1988, Kottke introduced a study in the area of municipal and regional tourism planning which used a DEA similar linear programming model. This model was designed to provide planning officials with a method for estimating the potential economic impact of tourism growth on a community. Although the model could be successfully applied, it did not deliver any information about the efficiency of an operational unit.

Hruschka (1986) and Banker and Morey (1986b) were the first to apply DEA to the hospitality industry, specifically to the restaurant sector. Hruschka analysed the panel database of the Austrian Society for Applied Research in Tourism, the same database which is the basis for the present study. He proposed and studied a form of efficiency measurement on an aggregated, rather than on an individual company level. In his study he performed DEA for ten different restaurant groups which uncovered differences in efficiency among those groups. Ten years later, another application of DEA for the measurement of restaurant productivity was introduced by Andersson (1996).

Bell and Morey (1994, 1995) introduced DEA for the use of benchmarking to discover best practice solutions in corporate travel management. They suggest allocative data envelopment analysis, an extension to the basic DEA, as the main benchmarking tool and examine this technique in an illustrative comparison of 31 travel departments. Bell and Morey's contribution to the benchmarking literature is particularly important, as they were the first to highlight the strengths of DEA for the selection of comparison partners.

There is a growing number of DEA applications in the hotel sector (Morey and Dittman, 1995, 1997; Johns *et al.*, 1997; Avkiran, 1999; Anderson *et al.*, 2000; Tarim *et al.*, 2000). Morey and Dittman (1995) gathered input–output data for 54 hotels of a national chain in the US. Using data for each individual hotel in the sample they applied DEA to generate a 'composite efficient benchmark general manager' which acts as a scorecard for the hotel under review. In fact, Morey and Dittman were the first to mention the use of DEA for developing instruments for use in the relatively new scorecard management philosophy. They also stressed the usefulness of DEA technique in the evaluation of franchising relationships, which is commonly used in the hotel sector. They note:

> In the hotel world, where the owner (franchisee), the manager, and the flag (franchiser) are often different entities, the owner needs to be able to evaluate the manager, and the flag needs to be able to provide service to the owner. One way for the flag to provide value to independent owners or to management companies that manage under many different flags would be to furnish an analysis like the one we recommend. (Morey and Dittman, 1995: 35)

From a macro perspective, they found that managers in the hotel industry were operating more efficiently than managers in other industries, which is in contradiction to findings by other authors who identified high levels of inefficiency caused by limited management skills in the industry (Baker and Riley, 1994). But Morey and Dittman argue that their findings are more consistent as large efficiency scores are typical for industries with high levels of performance and competition (Leibenstein, 1966).

Several studies appeared in the late 1990s. In a later study, Morey and Dittman (1997) apply DEA to selecting a hotel property. Their model combines

DEA and regression analysis. Their analysis approach maximizes the expected value of annualized profits given brand, design and operational choices (Morey and Dittman, 1997). Johns *et al.* (1997) used DEA to monitor and benchmark productivity in a chain of 15 hotels. Data for a 12-month period were used. Quarter results were compared with each other and with standard accounting data for the same period. They found DEA to be useful for diagnosing and identifying outstanding behaviour in terms of their measured productivity and gross profit. Similar to Johns *et al.* (1997), Avkiran (1999) used seasonal time series data for a small set of Australian hotel companies (23 units) in a case-study-like presentation.

More case studies were introduced recently. For example, Anderson *et al.* (2000) estimated efficiencies in the US hotel industry using data obtained from Ward's Business Directory of private and public companies. Their sample of 48 hotels captured data on total cost incurred by the hotel, the output produced and the input prices. Four output and six input measures were defined by the authors.

- Output measures:
  1. Total revenue generated from rooms
  2. Total revenue generated from gaming
  3. Total revenue generated from food and beverage
  4. Total other revenues
- Input measures:
  1. The average price of a room
  2. The average price of an employee
  3. The average price of food and beverage operations
  4. The average price of casino/gaming operations
  5. The average price of hotel/accommodation operations
  6. The average price of other expenses

Anderson *et al.* employed DEA to measure various forms of efficiency levels (overall, allocative, technical, pure technical and scale efficiency). Their findings, surprisingly, revealed that the hotel industry is highly inefficient with a mean overall efficiency measure of approximately 42%. In order to examine the source of allocative inefficiencies, the input mix of the more-efficient firms was compared with the input mix of the less-efficient firms. The results of this examination suggested that the efficient firms allocated more resources to food and beverage operations, whereas the less-efficient firms spend more on hotel operations and other expenses, employ too many workers and are too large in terms of the number of rooms. Overall, the number of implications and recommendations for hotel managers that could be drawn by the authors demonstrated the strengths of the DEA approach in efficiency measurement for the hotel industry.

Another case study measuring DEA efficiency in the hotel sector was reported by Tarim *et al.* (2000). In this study, the relative efficiency of four- and five-star hotels in Antalya (Turkey) were measured using the DEA approach; 21 hotels provided information on three output and three input measures.

- Output measures:
  1. Net profit
  2. Occupancy rate
  3. Customer loyalty
- Input measures:
  1. Investment costs
  2. Number of personnel
  3. Periodical administration expenses

Customer loyalty, an example of a non-financial performance indicator, was measured by the repeat-visitor ratio estimated by the hotel managers. A constrained DEA model formulation was used to consider the limited value range inherent in two of the three output variables (occupancy rate and customer loyality). By examining the DEA results the authors found that the four-star hotels in their sample were significantly more efficient than the five-star hotels. The main factors responsible for the efficiency evaluation were the repeat-visitor ratio and the net profit generated by the hotels. Similar to the study of Anderson *et al.* (2000), this study demonstrated the practicality of the DEA method and opened new possibilities to interpret the efficiency of different hotel categories in competitive terms.

Mathematical formulation of the general DEA (CCR model)

In DEA, the efficiency measure of a Decision Making Unit is defined by its position relative to the frontier of best performance established. In one of the early papers Charnes *et al.* (1981: 669) gave their formal definition of efficiency. They defined a firm as being 100% efficient when and only when:

1. None of its outputs can be increased without either
   (i) increasing one or more of its inputs, or
   (ii) decreasing some of its other outputs; and
2. None of its inputs can be decreased without either
   (i) decreasing one or more of its outputs, or
   (ii) increasing some of its other inputs.

This definition accords with the economist's concept of Pareto optimality and is free of the arbitrariness of any weighting or price imputation used in other approaches.[8] Charnes and Cooper (1985) argue that if there exists no other method for establishing a 'true' or theoretical model of efficiency, the previous definition must be adapted so that it refers to levels of efficiency relative to known levels attained elsewhere in similar circumstances. Hence, they state: '100% relative efficiency is attained by any unit only when comparisons with

8 A comprehensive description of this optimality concept is given by Koopmans (1951).

other relevant units do not provide evidence of inefficiency in the use of any input or output' (Charnes and Cooper, 1985).

Hence, by definition, DEA must allow one to identify, either for a given level of output the companies that achieved the lowest observed costs, or for a given level of costs the companies that achieved the highest observed output. For instance, for a given level of output Y the costs of each company X can be expressed as the sum of the minimum feasible costs plus an efficiency factor u reflecting the efficiency difference between the particular company and the 'best practice' companies, simply denoted as:

$$X = \min(\{X, Y\}) + u \tag{4.23}$$

Note the difference from Equation 4.1 for the simple regression model. DEA may be thought of as an alternative line-fitting algorithm that, instead of trying to fit a regression line through the centre of the data, floats a piecewise linear surface to rest on top of the observations (Seiford and Thrall, 1990: 8). Figure 4.7 recalls the formerly introduced regression example in Fig. 4.1, which will now be used to illustrate the DEA approach.

As before, the black dots plotted indicate companies with varying cost/output combinations $\{X,Y\}$. DEA labels companies such as P_1, P_2, P_3 and P_4 as efficient or 'best practice' companies because no other companies with the same output and lower costs can be identified. The line connecting all the efficient companies is the fitted DEA frontier, which envelops the data (hence the name Data Envelopment Analysis). DEA also labels companies such as P_5 as inefficient because, compared with its counterpart company P_3, it has the same level of output but higher costs. The distance P_3P_5 is a measure of the inefficiency of company P_5.

What about company P_6? Company P_6 has a different level of output from all the other companies and therefore cannot be compared with any of them. However, DEA compares it with the notional or artificially constructed company X_6* (appearing at the intercept between the dashed line through P_6 and the frontier in Fig. 4.7), which is a linear combination of companies P_2 and P_3. The companies P_2 and P_3, from which the notional company X_6* is created,

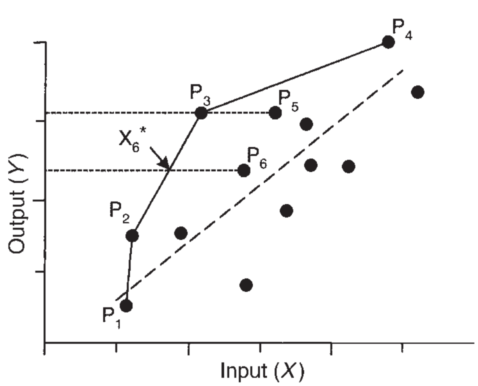


Fig. 4.7. Illustration of the DEA evaluation.

are said to be the 'peer group members' of P_6 (Bell and Morey, 1994). The distance P_6X_6* is then a measure of the efficiency of company P_6. Compared with its benchmark, X_6*, company P_6 is inefficient because it produces the same level of output but at higher cost.

Note that companies P_5 and P_6 are located above the average regression line and hence are identified as efficient companies by the linear regression method (see Fig. 4.1), but identified as inefficient by DEA.

One of the major achievements of Charnes *et al.* (1978) was that they generalized the DEA technique to handle simultaneously multiple cost drivers and multiple types of costs. In this generalized DEA approach results are almost never affected by the number of variables in the model. On the other hand, with regression analysis, as the number of explanatory factors increases and the sample size remains the same, there is a larger number of parameters to be estimated which reduces the available degrees of freedom, and therefore limits its use for estimation and hypothesis testing. Certainly, this is one of the reasons why DEA gained much popularity since the early 1980s.

In such a multiple DEA setting with *m* input and *s* output factors, an efficiency measure for a company *o* can be stated as:

$$\frac{w_1 y_{1o} + w_2 y_{2o} + \ldots + w_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo}} = \frac{\sum_{j=1}^{s} w_j y_{jo}}{\sum_{i=1}^{m} v_i x_{io}} \tag{4.24}$$

where y_j is the amount of output j, x_i is the amount of input i, w_j is a weight assigned to output j, and v_i is a weight assigned to input i. This efficiency measure is sometimes also referred to as 'virtual output/input ratio'. In DEA, the weights are derived from the data instead of being fixed in advance. For example, suppose a situation with two input and two output factors where somebody wants to compare the performance of two companies:

$$\frac{w_1 y_{11} + w_2 y_{21}}{v_1 x_{11} + v_2 x_{21}} = \frac{w_1 y_{12} + w_2 y_{22}}{v_1 x_{11} + v_2 x_{21}} \tag{4.25}$$

For the efficiency evaluation of company 1, it is simply necessary to determine a set of weights that will improve its standing relative to its competitor. An easier task than solving Equation 4.25 is to solve the equivalent problem:

$$\max \frac{w_1 y_{11} + w_2 y_{21}}{v_1 x_{11} + v_2 x_{21}} \tag{4.26}$$

subject to

$$\frac{w_1 y_{11} + w_2 y_{21}}{v_1 x_{11} + v_2 x_{21}} \leq 1$$

and

$$\frac{w_1 y_{12} + w_2 y_{22}}{v_1 x_{11} + v_2 x_{21}} \leq 1$$

with $w_1, w_2, v_1, v_2 > 0$. The weights are the maximum possible value of the ratio for the company under evaluation to the value of the ratio for the other company being constrained to a maximum of 1.

This can be generalized to a larger problem, comparing the performances of n companies. The equivalent efficiency measure for unit o is:

$$\max e_o = \frac{\sum_{j=1}^{s} w_j y_{jo}}{\sum_{i=1}^{r} v_i x_{io}} \tag{4.27}$$

subject to:

$$\frac{\sum_{j=1}^{s} w_j y_{jm}}{\sum_{i=1}^{r} v_i x_{im}} \leq 1 \qquad m = 1, \ldots, n$$

$$w_j \geq 0 \qquad j = 1, \ldots, s$$

$$v_i \geq 0 \qquad i = 1, \ldots, r$$

where the values of the ratio for unit o are less than 1, the subset of units whose ratio value is equal to 1 is the peer group for unit o.

Charnes *et al.* (1978) have proved that this fractional programming problem to obtain values for the input and output weights can be solved by a linear program (LP). For this purpose they introduce an additional constraint by setting the weighted sum of inputs to unity (= 1). This is possible since the objective is to maximize the denominator of the function in Equation 4.27. The problem can then be expressed as the following LP:

$$\max e_o = \sum_{j=1}^{s} w_j y_{jo} \tag{4.28}$$

subject to:

$$\sum_{i=1}^{r} v_i x_{im} - \sum_{j=1}^{s} w_j y_{jm} \geq 0 \qquad m = 1, \ldots, n$$

$$\sum_{i=1}^{r} v_i x_{io} = 1$$

$$w_j \geq 0 \qquad j = 1, \ldots, s$$

$$v_i \geq 0 \qquad i = 1, \ldots, r$$

Every LP has both a primal and a dual formulation which have identical solutions. It is not the intention to cover the theory of LP within this text since there are many standard textbooks on this subject (e.g. for a theoretical discussion see Dantzig (1963) and Dantzig and Thapa (1997); for applications in management science see Anderson *et al.* (1997); see also Cooper *et al.* (2000) for a comprehensive introduction to DEA models). The dual formulation gives by far the clearest interpretation of what is happening in terms of input minimization and output maximization. In this respect, Charnes *et al.*'s (1978) choice of introducing the primal formulation was an unfortunate decision. It is a source of constant confusion among practitioners and researchers who work with the primal to think that input minimization is achieved through the formulation that maximizes the objective function, whereas output maximization is achieved through the formulation that minimizes the objective function.

The dual formulation of Equation 4.28 is:

$$\min f_o \tag{4.29}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{im} \leq f_o x_{io} \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} x_{jm} \geq y_{jo} \qquad j = 1, \ldots, s$$

The model can be interpreted as follows:

> For company o, find the minimum proportion f_o which allows a weighted combination (i.e. the λ_{om}) of the performance of all companies to be found such that for each input, the weighted combination of input does not exceed the proportion f_o of the input of company o and for each output, the weighted combination of output is at least as great as that of company o. (Norman and Stoker, 1991: 237)

It is obvious that the minimum value of f_o will never be greater than 1 since a value of 1 can always be obtained via the performance of company o itself by choosing the weighted combination given by $\lambda_{oo} = 1$ and all other $\lambda_{om} = 0$.

A minimum value of f_o less than 1 determines the existence of a weighted combination of the actual performance of other companies such that no output of company o exceeds the corresponding output of the weighted combination and, at the same time, it is possible to reduce all of the inputs of company o to the proportion f_o of its existing value without any input falling below the corresponding input of the weighted combination. In other words, the weighted combination produces at least as much output in every instance for less input and f_o is a measure of how much all of the inputs of company o can be reduced in the same proportion to produce a performance in line with the weighted combination. Efficiency of companies can therefore be interpreted in the sense that no other company or weighted combination of companies can perform as well in every instance across the complete range of inputs and outputs.

In the LP formulation 4.29, it is clear from the mathematics that the inputs of company o can be reduced to the proportion f_o of their existing value at which point some of the input and output constraints are satisfied as equalities. To reduce all inputs further would result in violation of those constraints. However, equality may not be achieved in all constraints, in which case company o will still have some inputs greater than the inputs of the weighted combination or some outputs less than the outputs of the weighted combination, indicating that additional improvements are possible in those aspects.

The magnitude of the additional possible improvements is simply the amount by which the relevant constraint fails to achieve equality. These amounts can be included in the LP formulation as s_i^-, $i = 1, \ldots, r$, and s_j^+, $j = 1, \ldots, s$, as follows:

$$\min f_o \tag{4.30}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{im} - f_o x_{io} + s_i^- = 0 \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jm} - s_j^+ = y_{jo} \qquad j = 1, \ldots, s$$

The s_i and s_j are called 'slack variables' which is standard LP terminology for additional variables introduced to convert inequality constraints into equality constraints. This is also the terminology used in DEA when referring to the additional improvements possible in specific inputs or outputs.

In their introductory paper on DEA, Charnes *et al.* (1978) introduced the simple example reproduced in Fig. 4.8. The example depicts a DEA for six companies each having a single output with the same value and two inputs with differing values as plotted on the graph. The main point of the example is that, for company P_6, there are two possible minima. The minimum value of f_6 is 1, but this can be achieved for either $\lambda_{63} = 1$ and all other λ_{6m} and slacks zero; or $\lambda_{66} = 1$, $s_2 = 1$, all other λ_{6m} and slacks zero. In other words, for P_6, best performance is exhibited either by P_3 or by P_6, but in the case of P_6 with a reduction of input 1 by one unit, P_6 would become coincident with P_3. Hence

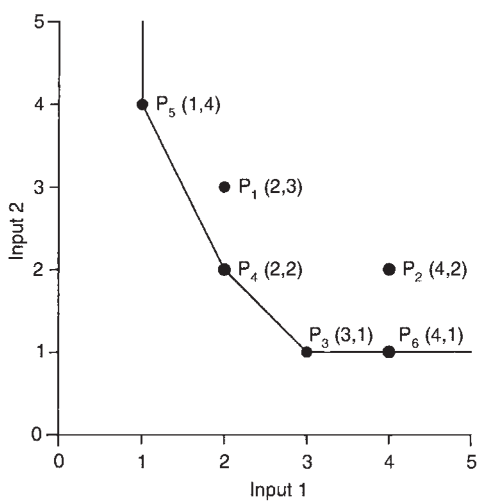


Fig. 4.8. Alternative minima in DEA models (Charnes *et al.*, 1978).

in practical terms there is no ambiguity, although in mathematical terms there are dual minima.

The problem arises from the radial Farrell measure because inputs are equiproportionately reduced. It is possible that there exists slack in some but not all of the inputs and/or outputs even after Farrell efficiency is achieved. Charnes *et al.* (1978) suggested removal of this mathematical ambiguity by amending the objective function to maximize the slack values too, but in a manner which did not impair the minimization of f. This resulted in the following amended objective function which has become the form usually quoted when the DEA dual formulation is presented:

$$\min\left\{ f_o - \delta\left(\sum_{i=1}^{r} s_i^- + \sum_{j=1}^{s} s_j^+ \right)\right\} \tag{4.31}$$

where δ is the non-Archimedean infinitesimal, usually a very small number.[9]

It is also possible to invert the efficiency measure to obtain an equivalent result with the efficiency measure becoming the reciprocal of its former value. Hence the problem can be transformed from 'Given the outputs, by how much can we reduce the inputs?' to 'Given the inputs, by how much can we increase the outputs?', and perhaps not surprisingly, the answer to one is the reciprocal of the other. However, although the two formulations are equivalent in terms of measuring efficiency, a different value for slack is obtained depending on which of these formulations is used. For a complete formulation of the reciprocal equations of the LP problem see Charnes *et al.* (1994b) or Norman and Stoker (1991).

SCALE ASSUMPTIONS A distinguishing feature of DEA is that it can be used to handle cardinal and non-cardinal data. Particularly, Banker and Morey (1986a) showed how categorical variables can be incorporated in DEA models. Their model was later refined by Kamakura (1988) and Rousseau and Semple (1993). Finally, Cook *et al.* (1993) illustrate how to use qualitative information, measured in the form of ordinal data, in DEA applications.

UNCONTROLLABLE INPUTS The effect of uncontrollable inputs can be incorporated into DEA when reductions cannot be achieved. This development is included in a paper by Charnes and Cooper (1985) where they make use of the terms discretionary and non-discretionary inputs – equivalent to controllable and uncontrollable inputs. In the modified weighted sum ratio the controllable input remains in the numerator with the uncontrollable input being incorporated into the denominator. The equivalent efficiency measure for company *o*

9 In order to maintain the units-invariance property of the objective function values, it is important that δ be very small. Numerical difficulties can arise, however, in computation, with the selection of an arbitrarily small number (e.g. 10^{-6}) to approximate δ, as many standard LP software packages do (Ali, 1994). The correct algorithmic implementation requires a two-stage (preemptive) approach.

with s outputs denoted by y_j, with $j = 1, \ldots, s$, r controllable inputs denoted by x_i, $i = 1, \ldots, r$, and t uncontrollable inputs denoted by z_k, $k = 1, \ldots, t$ formulated as a fractional optimization problem is:

$$\max e_o = \frac{\sum_{j=1}^{s} w_j y_{jo} - \sum_{k=1}^{t} u_k z_{ko}}{\sum_{i=1}^{r} v_i x_{io}} \tag{4.32}$$

subject to:

$$\frac{\sum_{j=1}^{s} w_j y_{jm} - \sum_{k=1}^{t} u_k z_{km}}{\sum_{i=1}^{r} v_i x_{im}} \leq 1 \qquad m = 1, \ldots, n$$

$$w_j \geq 0 \qquad j = 1, \ldots, s$$

$$v_i \geq 0 \qquad i = 1, \ldots, r$$

$$u_k \geq 0 \qquad k = 1, \ldots, t$$

As before, this can be rearranged as an LP as follows:

$$\max e_o = \sum_{j=1}^{s} w_j y_{jo} - \sum_{k=1}^{t} u_k z_{ko} \tag{4.33}$$

subject to:

$$\sum_{i=1}^{r} v_i x_{im} + \sum_{k=1}^{t} u_k z_{km} - \sum_{j=1}^{s} w_j y_{jm} \geq 0 \qquad m = 1, \ldots, n$$

$$\sum_{i=1}^{r} v_i x_{io} = 1$$

$$w_j \geq 0 \qquad j = 1, \ldots, s$$

$$v_i \geq 0 \qquad i = 1, \ldots, r$$

The dual of this LP is:

$$\min f_0 \tag{4.34}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{im} \leq f_o x_{io} \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} z_{km} \leq z_{ko} \qquad k = 1, \ldots, t$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jm} \geq y_{jo} \qquad j = 1, \ldots, s$$

Again, an excellent interpretation of the above model is given by Norman and Stoker:

> For the company *o*, find the minimum proportion f_o which allows a weighted combination (i.e. the λ_{om}) of the performance of all companies to be found such that for each controllable input, the weighted combination of input does not exceed the proportion f_o of the input of company *o*, for each uncontrollable input, the weighted combination of input does not exceed that of company *o*, and for each output, the weighted combination of output is at least as great as that of company *o*. (Norman and Stoker, 1991: 244)

As for Equation 4.29, the dual provides a direct interpretation where the weighted combination produces at least as much output in every instance for less input and f_o is a measure of how much all of the controllable inputs can be reduced in the same proportion to produce a performance in line with the weighted combination. The only difference between this and formulation 4.29 is that f_o is only applied to the controllable inputs which is consistent with the definitions of controllable and uncontrollable. In fact, Equation 4.29 is simply a special case of Equation 4.34 where $t = 0$.

The reciprocal formulation of the DEA model with controllable and uncontrollable inputs is of little value because adapting the above formulation for output maximization would be factoring the uncontrollable inputs as well as the outputs. However, lacking control of uncontrollable inputs is undesirable and so, in this case, the reciprocal formulation is inappropriate.

VARIABLE RETURNS TO SCALE (BCC MODEL) In an environment with increasing or decreasing returns to scale DEA results will vary depending on the orientation of the model. This is depicted in Fig. 4.9(a) where decreasing returns to scale technology is represented by $f(x)$, and an inefficient firm operating at point P.

The input-oriented measure of technical efficiency would be equal to the ratio $\overline{AB} \cdot \overline{AP}^{-1}$, while the output-oriented measure of technical efficiency would be $\overline{CP} \cdot \overline{CD}^{-1}$. The output- and input-oriented measures will only provide

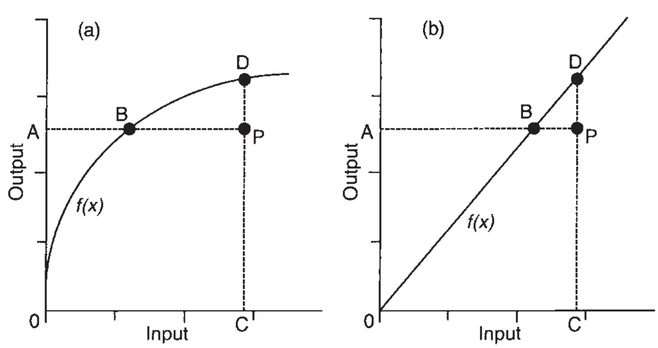


Fig. 4.9. Returns to scale effect on DEA models. (a) Decreasing; (b) constant.

equivalent measures of technical efficiency when constant returns to scale exist, but will be unequal when increasing or decreasing returns to scale are present (Färe and Lovell, 1978). The constant returns to scale case is depicted in Fig. 4.9(b) where $\overline{AB} \cdot \overline{AP}^{-1} = \overline{CP} \cdot \overline{CD}^{-1}$ for any inefficient point P.

The Charnes, Cooper and Rhodes (CCR) model described in the previous section, is only appropriate when all companies are operating at an optimal scale. Imperfect competition, constraints on finance, etc. may cause a company to not be operating at optimal scale. Banker *et al*. (1984) suggest an extension of the CCR-DEA model to account for variable returns to scale. Based on the authors' initials, it is referred to as the BCC-DEA model. The use of the CCR model when not all companies are operating at the optimal scale will result in measures for technical efficiency that are confounded by efficiencies of scale. The BCC model permits the calculation of technical efficiency devoid of these variable returns to scale effects.

Banker *et al*. (1984) modify the CCR model to account for variable returns to scale by adding the convexity constraint $\sum_{m=1}^{n} \lambda_{om} = 1$ to Equation 4.29 to provide:

$$\min f_o \tag{4.35}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{im} \leq f_o x_{io} \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jm} \geq y_{io} \qquad j = 1, \ldots, s$$

$$\sum_{m=1}^{n} \lambda_{om} = 1$$

This approach forms a convex hull of intersecting planes which envelops the data points more tightly than the constant returns to scale conical hull and thus provides technical efficiency scores which are greater than or equal to those obtained using the CCR model. The variable returns to scale model provided by Banker *et al*. (1984) has been the most commonly used DEA model in the 1990s.

When conducting both CCR- and BCC-DEA on the same data, the technical scores can be decomposed into two components, one due to scale inefficiency and one due to 'pure' technical inefficiency (Banker, 1984). A difference in the two technical efficiency scores for a particular company indicates that the company has scale inefficiency, and the scale inefficiency can be calculated from the difference between the variable returns to scale technical efficiency score and the constant returns to scale efficiency score (Banker and Thrall, 1992). Figure 4.10 which is a one-input, one-output example illustrates this (Coelli, 1996). It shows DEA results under (i) constant

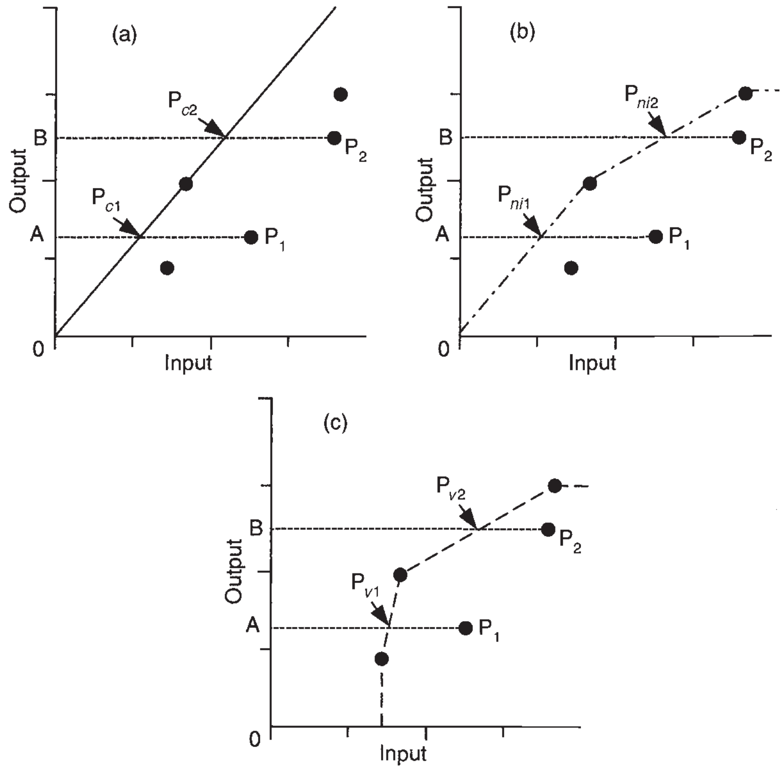


Fig. 4.10. Various returns to scale specifications in DEA. (a) Constant; (b) non-increasing; (c) variable.

returns to scale, (ii) non-increasing returns to scale and (iii) variable returns to scale assumptions applied to the same data set. For example, under the constant returns to scale assumption the input-oriented technical inefficiency of point P_1 is the distance $\overline{P_1P_{c1}}$, whereas under the variable returns to scale model the technical inefficiency would only be $\overline{P_1P_{v1}}$. The difference between these two, $\overline{P_{c1}P_{v1}}$, is put down to scale inefficiency.

One shortcoming of measuring scale inefficiency by comparing results from CCR and BCC models is that the value does not indicate whether the company is operating in an area of increasing or of decreasing returns to scale. This may be determined by running an additional DEA problem with non-increasing returns to scale imposed (Coelli, 1996). This can be done by altering the DEA model in Equation 4.35 by substituting the $\sum_{m=1}^{n} \lambda_{om} = 1$ restriction with $\sum_{m=1}^{n} \lambda_{om} \leq 1$.

The non-increasing returns to scale frontier is plotted in Fig. 4.10(b). The nature of the scale inefficiencies (i.e. due to increasing or decreasing returns to scale) for a particular company can be determined by seeing whether the non-increasing returns to scale technical efficiency score is equal to the variable returns to scale efficiency score. If they are unequal (as will be the case for the point P_1 in Fig. 4.10), then increasing returns to scale exist for that company. If they are equal (as is the case for point P_2 in Fig. 4.10) then decreasing returns to scale apply. There are several alternatives of how to determine returns to scale in DEA; for more recent discussions see Banker *et al.* (1996) and Seiford and Zhu (1999).

Given the importance of uncontrollable variables it should be noted that, in analogy to the CCR model, the BCC can also be reformulated to handle uncontrollable variables.

4.2.2 Stochastic frontier models

DEA is popular in literature primarily because it does not require any assumption about functional form and can readily handle multiple inputs and outputs. However, DEA has been criticized strongly in the literature as having serious potential statistical shortcomings. One of the major objections to the DEA approach has been that the frontier itself is determined by extreme observations of the data set. Thus the definition of the frontier could be sensitive to errors or inconsistencies in the data (Scheel and Scholtes, 1998). In general, DEA methodologies assume that there are no random fluctuations from the efficient frontier. Therefore, DEA is sensitive to the input/output specification of the model as it can only measure relative efficiency levels. Hence, if one firm in the sample is much more or much less efficient than the average firm in the sample, the DEA will produce large inefficiency measures. Additionally, DEA is a non-parametric technique that does not allow for random deviations from the efficient frontier. Therefore, any deviation from the frontier is deemed inefficiency. This type of classification tends to overestimate inefficiencies.

Considerable research effort has been expended in the development of more sophisticated methods of evaluating the efficiency of a company in relation to the other companies in its grouping. The stochastic frontier production function was independently proposed by Aigner *et al.* (1977) and Meeusen and van den Broek (1977). The original specification involves a production function specified for cross-sectional data which incorporates a two-part error term, one to account for random effects and another to account for technical inefficiency. This model can be expressed in the following form:

$$y_i = x_i\beta + (v_i - u_i) \qquad i = 1, \ldots, N \tag{4.36}$$

where

$$u = |U| \text{ and } U \sim N\left[0, \sigma_u^2\right]; v \sim N\left[0, \sigma_v^2\right]$$

The production (or the logarithm of the production) of the ith company is denoted by y_i. The input quantities of the ith company is a $k \times 1$ vector of x_i, and β is a vector of unknown parameters. The components of the disturbance are assumed to be independent. U are non-negative random variables which account for technical inefficiency in production and are often assumed to be normally distributed. Similarly, v will account for the random effects in the model.

There are many variations on this that have appeared in literature. All techniques attempt to define variations from an efficient frontier using alternative assumptions regarding the probability distributions of the technical efficiency and random error. The specification in Equation 4.36 has been altered and extended in a number of ways. These extensions include the specification of more general distributional assumptions for u, such as the truncated normal, normal–half normal, or normal–exponential model, the consideration of panel data and time-varying technical efficiencies (Cornwell *et al*., 1990; Kumbhakar, 1990), the extension of the methodology to cost functions, etc. A number of comprehensive reviews of this literature are available, such as Førsund *et al*. (1980), Schmidt (1986), Bauer (1990), Ferrier and Lovell (1990), Greene (1993) and Simeone and Li (1997).

Similar to the non-parametric DEA for frontier estimation, the stochastic frontier approach can provide individual firm efficiency estimates. Hence, the technique allows managers and researchers, by comparing, to identify companies (or departments) that are relatively efficient, to determine the magnitude of firms' inefficiencies and to suggest alternative paths to reduce inefficiencies.

Although several studies compare efficient frontier techniques, the literature is not clear on which approach is superior (Anderson *et al*., 1999). Because of its ability to include the possibility of deviations from the efficient frontier to be a function of random error, econometricians generally favour this approach (Simeone and Li, 1997). What is clear from the literature is that the choice of technique can dramatically influence the results (Berger *et al*., 1993; Reinhard *et al*., 2000; Worthington, 2000). For example, using similar data for the banking market, Berger *et al*. (1993) found a very inefficient market using DEA, while the same market was shown to be relatively efficient using a stochastic frontier approach.

However, there are also several researchers that find DEA superior when compared to stochastic frontier or other statistical approaches. For example, Bjurek *et al*. (1990) performed parametric and non-parametric frontier analysis on a data set of 400 Swedish social insurance companies and found the differences between the two methodologies 'surprisingly small' (p. 227). Banker *et al*. (1993) compare the production frontier estimation by corrected ordinary least squares (COLS) and DEA for varying sample sizes between 25 and 200 and find that DEA outperforms COLS for small sample sizes and both methods are equally poor for high measurement errors. Also more recently, Ruggiero found that stochastic frontier analysis cannot produce an accurate decomposition of the total error into inefficiency and noise components. His results suggest that at best, the stochastic frontier is only as good as the deterministic model (Ruggiero, 1999: 562).

Similar results are found by Gstach and Resti who performed several comprehensive tests on the robustness of DEA results with simulated data including Monte Carlo analysis and bootstrapping approximations (Gstach, 1993a,b, 1994, 1996; Resti, 2000). Gstach also finds DEA superior in studies of small sample size and he considers DEA to be a serious competitor to stochastic production frontier estimation (Gstach, 1994: 8). For DEA applications in very noisy environments, he proposes a new two-stage approach where, in stage one, DEA estimates a pseudo frontier and, in stage two, a maximum-likelihood technique is used to estimate a value by which this pseudo-frontier must be shifted to get the true production frontier (the DEA+ model is discussed in Gstach, 1996). Hence, for the simple purpose of robustness, it seems important to examine efficiency using multiple techniques.

To the knowledge of this author there is only one study on stochastic frontier analysis with an application to the hospitality industry. Anderson *et al.* (1999) use this technique to estimate managerial efficiency levels in the hotel industry. They used data for 48 US hotels they obtained either from annual reports or from the Internet. The companies represent a wide cross-section of hotels from varied regions of the country and were very heterogeneous in terms of size. In their study they use a translog cost function with six input prices and one output. They define output for the hotel as the total revenue generated from rooms, gaming, food and beverage, and other revenues; inputs are defined as those items that the hotel uses to produce total revenue, which are the number of full-time equivalent employees, the number of rooms, total gaming-related expenses, total food and beverage expenses, and other expenses. Due to data constraints, the price of an employee was proxied by total hotel revenues divided by the product of the number of rooms, the occupancy rate and days per year. The price of gaming, food and beverages, and other expenses were proxied by estimating each as a percentage of total revenue.

Although the use of price estimates is certainly a limitation in their study, the translog cost function specified by the authors fits the data well. Similar to Morey and Dittman (1995) they found efficiency scores to be very high relative to what is found in other industries (mean: 89.4%; median: 89.6%). The slightly higher efficiency scores found in their study were expected by the authors because of the differences in the estimation techniques employed.

Recently, Worthington (2000) considered the underlying reasons for the very different levels and rankings of efficiency that the author experienced during a study on the performance of Australian local governments. After comparing the pros and cons of both techniques the author concluded that they are complementary analytical tools for different purposes: DEA reflecting relative efficiency suitable for benchmarking and the stochastic frontier measuring absolute efficiency.

Chapter 5

The Austrian Hotel and Restaurant Panel

This chapter introduces the database used in subsequent experimental studies. First, information is provided on the characteristics of the Austrian lodging industry and its development during the period under investigation. Next, the methodology used for collecting the business data is described. This includes describing the subsample selected for the further analysis. Finally, the prototype version of the decision support system for financial benchmarking is introduced and discussed.

The development of the Austrian Hotel and Restaurant Panel (AHRP) was an initiative of the Austrian Professional Hotel Association and the Austrian Professional Restaurant Association of the Federal Chamber of Commerce which started in the late 1960s. These associations have realized that Austrian small and medium-sized enterprises cannot cope with a deficiency of comparative financial information on their own. Since 1968 these two organizations have financed an annual survey of Austrian hotel and restaurants that is performed by the Austrian Society for Applied Tourism Research (ASART).

The development of the AHRP during the 1970s and 1980s is a credit to many people working for ASART, but it was especially enriched by the effort of Erwin Plank, who established the uniform system which was appreciated by the Austrian hotel and restaurant industry and became the basis for the database design (Plank, 1973).

Analysis in this book is based on data spanning a period of 7 years from 1991 to 1997. In 1991 the author became involved with the AHRP project. Since then, the author explored many aspects of performance issues in the hotel and restaurant sector which has resulted in several industry reports (Wöber, 1994–2000, 1996; Mazanec *et al.*, 1996) and research notes and conference contributions (Wöber, 1993, 1999, 2000).

5.1 The Austrian Hotel Industry Between 1991 and 1997

In 1998 Austria attracted 25 million tourists and recorded 111.1 million bednights.[1] Austria is one of the major tourist destinations in the world with approximately 3% market share of world tourism. Tourism plays a vital role in the Austrian economy as it is an extremely valuable source of income, employment and foreign currency earnings. However, since 1992 Austrian tourism has seen a continuous decline both in terms of the total number of arrivals and bednights. As shown in Fig. 5.1, since 1992, when Austrian tourism recorded its highest level with 130.4 million bednights, the number of bednights has declined by 14.2%.

Austria has been severely affected by the general relative weakening of currencies, in particular the US dollar. Holidays in Austria have become less affordable. In Europe, the competitive devaluation of the Italian, Spanish and UK currencies during the summer of 1992 strongly affected tourist flow from these important origin markets. A second phenomenon has been increased competition as a result of the emergence of new destinations in Europe like the Czech Republic and Hungary. Other countries such as Turkey and Spain have consolidated their positions thanks to massive increases in their accommodation capacities. Other explanations cite that Austria is perceived as an

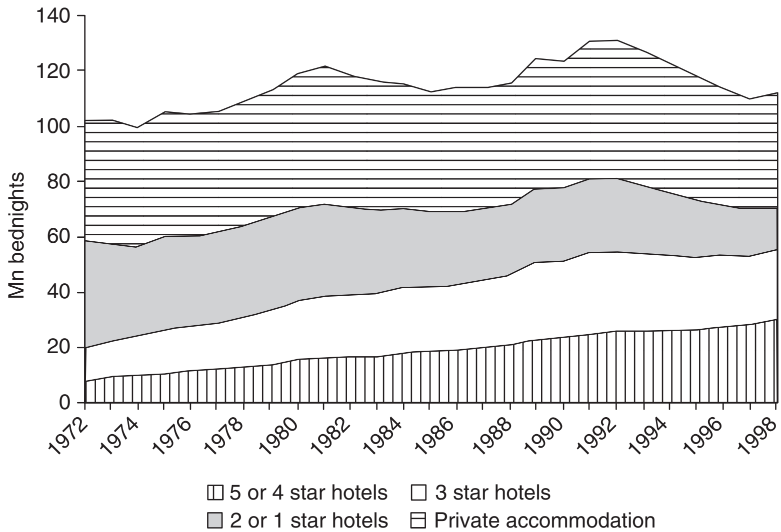


Fig. 5.1. Austria's bednights statistics between 1972 and 1998.

1 Bednights refers to the number of nights per person, e.g. two people staying for one night in one room (even in one bed) generate two bednights.

old-fashioned country, or that lower and moderate-priced accommodation suffers from the very high expectations of quality of service by foreign tourists not being met.

The hotel sector within the AHRP was chosen for this study for a variety of reasons. When this study was launched the hotel market in Austria was undergoing some upheaval due to a drop in tourism demand in the country. Austria possesses relatively well-developed statistics on the amount of accommodation generated by various accommodation suppliers. Change and good statistics being available made the market intriguing and promised some insights into how the various performance measurement techniques will reflect the recession in tourism demand, which was accompanied by heavy investments necessary to accomplish the increased level of quality demanded by the customers.

The quality of accommodation in Austria improved dramatically between 1991 and 1998. Two main reasons explain this trend. First, the federal government pursued a policy that supported the upgrading of facilities. Second, for many years there has been a general trend towards shorter holidays, with higher quality services, especially for city tourism.[2]

According to regular publications of the Austrian Statistical Office,[3] the number of accommodation providers increased continuously until 1992. Then, due to the recession in tourism demand, a tremendous decline in the number of accommodation providers was observed. The number of classified commercial accommodation establishments[4] decreased from 18,975 in 1992 to 17,759 in 1998, but the number of four- and five-star hotels showed a continuous increase.[5] Their market share increased from 18.6% of bednights in 1991 to 26.0% in 1998. The market share of three-star hotels also slightly increased, from 22.5% in 1991 to 23.2% in 1998.

The total number of available beds in classified accommodation establishments reached 584,889 in 1998. This is a decrease of 10.6% compared to 1991. It primarily reflects the elimination of low quality accommodations.

Most hotels in Austria are run on an independent basis or belong to local hotel chains. During the last decade, though, major international chains have increased their presence in Austria, with private accommodation[6] still representing 54% of the total number of beds available (695,990 beds) and accounting for 37% of total overnights in 1998.

---

[2] The average duration of stay of tourists in Austria dropped from 6.6 days in 1972 to 4.5 days in 1998.

[3] See also tourmis.wu-wien.ac.at

[4] Accommodation establishments that are registered members of the Austrian Chamber of Commerce and who are classified by this organization's hotel classification system.

[5] Tourism in Austria has two distinct seasonal peaks: winter, from late November–early December to February–March; and summer, concentrated from June to the end of August. The data presented here refer to summer capacities. The winter capacities developed similarly and are regularly analysed by Wöber (1994–2000).

[6] Primarily bed and breakfast providers, but also youth hostels, camping sites, holiday apartments and cure centres.

5.2 Survey and Database Design

For the AHRP, financial information is collected annually on a confidential basis from between 1000 and 1300 hotels and restaurants. The data are collected either from interested companies that voluntarily participated in this project (15.0%), consultancy companies (5.5%) and cooperating industry organizations which are Österreichische Hotel- und Tourismusbank (55.6%) and BÜRGES Förderungsbank des BmwA (23.9%).[7] The sample size varies between 2 and 3% of all hotels, restaurants and similar accommodation and/or food and beverage (F&B) providers in Austria, which is similar to other studies in this field.

During the seven years under observation, the number of participating hotels and restaurants was relatively stable with an exception in 1995 when a technical problem by one of the cooperating industry organizations caused a considerable drop in the number of participating companies (see Fig. 5.2).

The data comprise information from a company's balance sheet and the profit and loss statement, and information obtained from an additional questionnaire. The detail of information in the accounting part of the survey follows the most common definitions in the Austrian hotel and restaurant accounting system.

Performance analysis solely based on accounting information does not reflect many aspects of operational productivity and neglects important differences between various forms of businesses. For instance, accounting data usually do not capture information about the number of employees and the

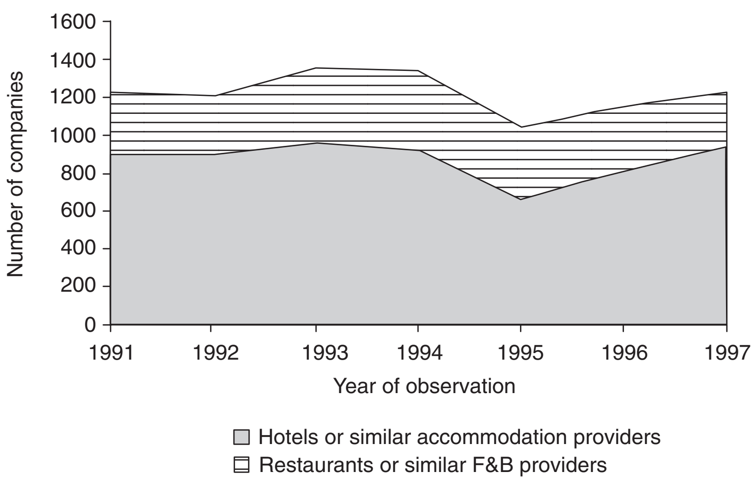


Fig. 5.2. Number of companies in the AHRP.

7 Percentages represent average share of companies provided to the AHRP project by the different parties.

distribution of the personal resources in the different cost centres. Especially in the hotel sector, financial reports do not indicate the number of overnights generated during the fiscal year, nor do they give information about the available (maximum) capacities. Hence, even the simple calculation of more valuable ratios of productivity requires additional information on certain business characteristics.

In AHRP, the non-financial data are obtained using an additional questionnaire that is filled out either by the general manager himself or by the tax consultant in charge of the accounting issues. Information captured by questionnaire includes the type of accommodation or food and beverage supply, average annual capacity, number of employees, number of days of operation, number of overnights, the geographical region where the business is located and the form of ownership. A complete listing of variables in the data set for each individual company is presented in Table 5.1 for the balance sheet, in Table 5.2 for the gain and loss statement and in Table 5.3 for the additional

Table 5.1. Balance sheet data in AHRP.

No.	Variables
B211	Real assets
B212	Technical equipment and machines
B213	Other fixed assets
B210	Fixed assets
B221	Raw material and goods
B222	Receivables
B223	Cash and bank deposits
B224	Other current assets
B225	Accruals
B220	Current assets
B251	Equity capital (capital stock and earned surplus)
B252	Appropriated retained earnings
B253	Other equity capital
B250	Equity capital
B261	Long-term liabilities
B262	Current liabilities
B263	Deferrals
B264	Reserve for uncertain liabilities or anticipated losses
B260	Total liabilities
B240	Total assets

Table 5.2. Gain and loss statement in AHRP.

No.	Variables
B311	Food revenues
B312	Beverage revenues
B314	Accommodation revenues
B310	Main revenues
B320	Sideline revenues
B510	Total sales revenues
B330	Other revenues
B341	Exceptional revenues
B342	External revenues
B351	Liquidations of reserves and provisions
B352	Other revenues from financial transactions
B360	Total revenues
B411	Food costs
B412	Beverage costs
B415	Food and beverage (F&B) costs
B417	Other material costs
B410	Material-type costs
B420	Total payroll and related
B431	Energy costs
B432	Cleaning costs
B433	Maintenance costs
B434	Communication (phone fax and mail) costs
B435	Costs for promotion and representation
B436	Administration costs
B437	Taxes and contributions
B438	Rental and leasing costs
B439	Interest for loan capital
B441	Costs for material of low value
B447	Miscellaneous costs
B451	Total other payable costs
B473	Depreciation
B474	Endowment of reserves and provisions
B452	Not payable costs
B520	Operating costs
B460	Exceptional costs and sideline expenses
B480	Total costs
B500	Profit (before tax) or loss

Table 5.3. (Additional) Questionnaire data in AHRP.

No.	Variables (capacity)
S120	Number of overnights in the fiscal year
S100	Number of beds
S110	Number of rooms
S131	Number of seats in the restaurant
S132	Number of outside seats (garden or terrace)
S190	Total days of operation
S133	Days of operation (outside)
	Variables (employed staff)
S151	Number of employers
S152	Number of employees
S153	Number of trainees
S155	in the kitchen
S156	in restaurant service
S157	in accommodation service
S158	in administration
S150	Total number of staff
	Variables (other data not included in financial statements)
A100	Industry (full hotel, hotel garni, spa hotel, restaurant, inn, coffee-house, bar, buffet)
A110	Category (1–5 star system for hotels; category 1–2 for restaurants or similar)
A120	Location (village, small town, mid-size town, Vienna)
A130	Period of operation (seasonally or annually)
A140	Property conditions (perfect ownership, rented place)
S160	Value of fire insurance
S195	Gross investments in fiscal year to be capitalized
S170	Employer's remuneration
S175	Private deposit or withdrawal

information gathered from the questionnaire. The variable names displayed in the left column of these tables are used in equations and elsewhere as appropriate.

5.2.1 Limitations associated with accounting studies

Uniform accounting systems have existed for many years in several industries, particularly in the US. They are a set of principles and in some cases of accounting methods, which, when incorporated in the accounting system of the members of an industry, will result in the gathering of comparable financial figures. In the hospitality industry, the most well-known accounting standard is the Uniform Systems of Accounts, which was developed by the American

Hotel and Motel Association in 1961 and revised several times since then.[8] Chin *et al*. (1995) state that the Uniform Systems of Accounts facilitates performance measurement and comparison. Therefore, many international hotel chains have adapted to this norm, and in the US, 78% of all hotels, which have recently been reported by Kwansa and Schmidgall (1999). In Europe, due to characteristics of small and medium-sized entities, many countries decided to develop their own standards that are, however, in many aspects in accordance with the US rules.

An adaptation of the uniform accounting system was introduced in Austria that provided accounting guidelines considering the special needs of the Austrian lodging industry. However, although heavily promoted by industry representatives, many financial statements of Austrian hotels and restaurants still lack the desirable level of comparability. Some hotels and restaurants provide very detailed information whereas others are at a high level of aggregation.

To achieve a minimum level of comparability for the AHRP project, clear definitions are given to all data providers. These definitions not only include guidelines for virtually every possible figure in a firm's financial statement, but give answers to questions like 'Where are costs for water consumption to be included?' Care has been taken to deal with the classification of business operation, property, and the level of service (categorization scheme), the evaluation of seating and lodging capacities and the number of employees (assessment of half-time/part-time workers). The definitions are printed next to the data entry fields of the evaluation form and are also available at the AHRP Internet site at TourMIS illustrated and discussed in the next paragraph.

At the end of the data entry for a specific fiscal year, plausibility checks are performed. Based on the check, firms are excluded due to missing or questionable data. To address differences between reports by different firms, some of the general business data are aggregated or weighted in order to achieve acceptably comparable information in the data set. For instance, the total number of seats as an indicator for the F&B capacity of a hotel or a restaurant is calculated by the number of seats inside the restaurant/hotel building plus the number of seats available outside (garden or terrace) weighted by the number of days with outside operations divided by the total number of days in the fiscal year. A comprehensive list of all definitions and transformations performed on the data set is given by Wöber (1994–2000).

In order to allow more meaningful comparisons hotels and restaurants are classified according to their size, location (country or city), number of closing days and season of operation in 34 different strategic groups (industry sectors). Typologies of hotels based on criteria of ownership, size and opening pattern are commonly used in hospitality studies as they offer a more accurate picture of individual businesses (e.g. Sundgaard *et al*., 1998). For at least two reasons

[8] The Uniform Systems of Accounts for the Lodging Industry is now in its ninth revision (American Hotel and Motel Association, 1996). For more information see www.ei-ahma.org

the classification scheme has been kept unchanged since the beginning of the AHRP project. The first reason is to maintain continuity in the reporting. One important aspect, appreciated not only by interested (or participating) companies, but also by industry representatives, is the possibility of gaining a certain understanding of industry trends by comparing time series. Some investigation of time series of key ratios has been done for specific strategic groups within the industry. Secondly, budgetary restrictions on the project did not allow for the work of preparing new schemes, implementing these and revising editing and reporting programs to accommodate the changes. The definitions of the strategic groups outlined in the AHRP, and regularly analysed and reported by Wöber (1994–2000), are listed in Table 5.4.

5.3 The Basic AHRP Data Entry and Reporting System on the Web

One of the most challenging tasks in the 21st century is mastering the information explosion, caused in large part by the availability of increasingly sophisticated information-processing technology. From the business science

Table 5.4. Strategic groups in the AHRP.

Hotels or similar accommodation provider		Restaurants or similar F&B provider	
1	Cat. 5/4, Vienna, annually open	20	Restaurants, category 1, perfect ownership
2	Cat. 5/4, mid-size towns, annually open	21	Restaurants, category 1, rented place
3	Cat. 5/4, small towns, annually open	22	Restaurants, category 2, perfect ownership
4	Cat. 3, Vienna, annually open	23	Restaurants, category 2, rented place
5	Cat. 3, mid-size towns, annually open	24	Inns, category 1, perfect ownership, > 10 employees
6	Cat. 3, small towns, annually open	25	Inns, category 1, perfect ownership, ≤ 10 employees
7	Cat. 1/2, Vienna, annually open	26	Inns, category 1, rented place, > 10 employees
8	Cat. 1/2, mid-size towns, annually open	27	Inns, category 1, rented place, ≤ 10 employees
9	Cat. 1/2, small towns, annually open	28	Inns, category 2, perfect ownership
10	Cat. 5/4, seasonally open: summer and winter	29	Inns, category 2, rented place
11	Cat. 3, seasonally open: summer and winter	30	Coffee-house, category 1, perfect ownership
12	Cat. 1/2, seasonally open: summer and winter	31	Coffee-house, category 1, rented place
13	Spa hotels	32	Coffee-house, category 2, perfect ownership
14	Cat. 5/4, seasonally open: only summer	33	Coffee-house, category 2, rented place
15	Cat. 3, seasonally open: only summer	34	Bars and buffets
16	Cat. 1/2, seasonally open: only summer		
17	Cat. 5/4, seasonally open: only winter		
18	Cat. 3, seasonally open: only winter		
19	Cat. 1/2, seasonally open: only winter		

perspective researchers will have to allow themselves to be asked: Are managers getting the information they need? The decision process in the field of financial management involves the analysis of a large volume of data and information (Brignall and Ballantine, 1996: 23). Therefore, the need to access large databases and perform computations in real time is vital. This need has led researchers to implement decision support systems in various fields of financial management, such as financial planning, financial analysis and portfolio management. An extensive review of the literature on this subject was recently performed by Zopounidis *et al.* (1997). However, not a single reference of their 93 entries identifies a publication concerning a financial benchmarking system.

To foster meeting the increased information needs by Austrian hotel and restaurant entrepreneurs it was necessary for the AHRP to make use of the most promising communication instrument available today, the Internet. Since 1995, all findings in the annual report have been published on the Internet. In 1999, the Internet presentation was completely revised and an interactive system with direct database access was introduced by the author. This system was integrated into the Tourism Marketing Information System TourMIS[9] which is an Extranet[10] application developed by the Austrian National Tourist Office and the Austrian Society for Applied Research in Tourism with the objective of assisting tourism managers with market research and interactive decision support tools.

The AHRP site within TourMIS consists of a data entry section and a retrieval section. These are available for active and passive participants alike (Fig. 5.3). It should be noted that the system is accessible to all Austrian hotel and restaurant managers, including those who do not contribute to the survey.

5.3.1 Data entry

Each entrepreneur who wants to participate at the AHRP and is interested in using the Internet for data entry and reporting has to get an account within TourMIS. TourMIS offers all features of a user database that are necessary when assigning a personal account to each entrepreneur entering the system. These features include an online registration of users, an automatic assignment of a password via e-mail and the editing and deleting of his/her account.

After the entrepreneur's registration he enters his basic business data to the AHRP site and receives a company number which will be needed later for editing and in order to perform analyses. User identification, password and company number offer enough security for the participants. This is necessary

[9] See tourmis.wu-wien.ac.at
[10] For a comparison of Extranet applications in tourism see Marcussen (1998).

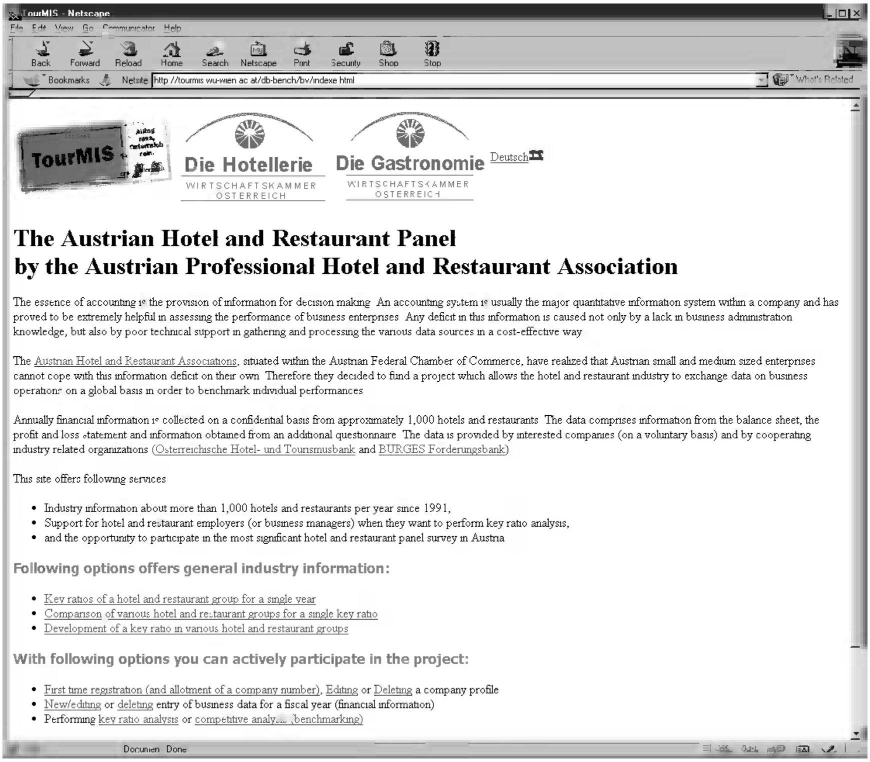


Fig. 5.3. Homepage of AHRP.

to guarantee the level of privacy someone expects when he hands in his financial information to AHRP.

Once the manager has entered his basic business data, which is static information and unlikely to change over the years (e.g. the company's address, location and type of operations), he/she can start entering financial data from the gain and loss statement and the balance sheet as well as the additional business data from the general questionnaire. Consistency checks during the storing phase (e.g. sum checks for the financial information) ensure that typing mistakes are avoided and the number of missing values are kept at a minimum. Financial data for multiple consecutive years can be entered into the system.

Although an entrepreneur may be deluged with printouts, reports and statistics from TourMIS, he or she may not necessarily be receiving the right kinds of information needed to plan and or manage his firm. For this reason, TourMIS has advanced protocol features in addition to the Web server's standard protocols, which allow the assessment of all inquiries and actions taken by the user. Assessment protocols linked to other protocols offered by TourMIS have been found to be a very valuable tool for learning the entrepreneurs' requirements and for the continuing development of the system (Wöber, 1998).

5.3.2 Data retrieval and analysis

Data entered by the user may pass all consistency checks during the data entry phase, but still be subject to errors. Although the compulsory identification and the built-in consistency checks support the avoidance of faked entries, this cannot be completely prevented due to the open system architecture. This characteristic of every interactive database application on the Web makes real-time analyses with a pooled data set difficult, as data for analysis may still be corrupt. In order to avoid this problem in the AHRP database the entrepreneur's business data are not directly saved to the general database, but first saved to a temporary database for all new entries. Periodically the system administrator checks each individual record in the temporary database and decides whether the record is a serious contribution to the project. Records accepted are moved to the general database and others are deleted. Whenever a manager of a company with an unproved data record requires the consideration of several companies for comparative analyses, the analyses are performed on the general database plus exactly one record (the record from the company under evaluation) from the temporary database. This procedure guarantees the integrity of the AHRP database and at the same time allows all conceivable procedures for the optimal selection of comparison partners.

5.3.3 The AHRP reporting functions

The system's basic retrieval and reporting functions offer the calculation of financial key ratios for each individual company plus the calculation of means, medians, maximum/minimum values and standard deviations for all strategic groups listed in Table 5.4. The reporting is organized by preformatted tables, which are generated on request by means of CGI scripts and an Xbase compatible database.[11]

The principal reports available within AHRP include:

1. The calculation of financial key ratios for one specific strategic group/industry sector;
2. The comparison of one specific financial key ratio for all strategic groups/industry sectors;
3. The development of a specific financial key ratio in all strategic groups/industry sectors;

[11] The program was developed in PERL and uses the Xbase – Perl module written by Pratap Pereira.

4. The development of all financial key ratios for a specific company over several consecutive fiscal years; and
5. The comparison of an individual company's financial key ratios with a specific strategic group/industry sector.

Reports 1 to 3 are general reports also available to non-participating hotel and restaurant managers or other interested parties accessing the system.[12] The tables generated by reports 4 and 5 require active participation by entering individual company data. Report 5 provides the simplest form of evaluation which in most hospitality industry reports is referred to as 'benchmarks' of industry performance (e.g. Bottomley, 1995; Horwath International, 1998; Pannell Kerr Foster, 1998).

5.3.4 Financial key ratios in the AHRP

When planning to open a new hotel or restaurant or measuring the performance of an existing one, certain essential relationships between various quantities are to be expected (Kotas and Kreul, 1987). One expects a certain relationship between the investment in assets and net profit, between net sales and net profit, between rooms available and rooms occupied, and so on. All such relationships may be expressed by ratios. There are numerous financial ratios that are commonly used by hotel managers to monitor business performance. In the AHRP study, 32 key financial ratios were calculated. These fall into four groups:

1. Ratios describing the operational performance in a specific fiscal year (derived from the gain and loss statement);
2. Ratios describing the financial situation of the company (derived from the balance sheet);
3. Ratios describing the employee's productivity; and
4. Ratios describing other characteristics of the company.

5.4 An Heuristic Procedure to Enhance the Basic AHRP System

One of the major problems encountered in this study was trading-off between the comparability of hotels or restaurants as a result of the classification into various industry sectors/strategic groups and the number of companies that could be used in a given analysis. The more detailed the grouping the fewer the companies in groups. For instance, requests concerning the database are sometimes very specific, e.g. 'What is the average occupancy rate for an 80-bed hotel-garni in Tyrol?' Such specific questions cannot usually be

[12] See tourmis.wu-wien.ac.at/db-bench/bv/indexe.html

processed since the database does not include a sufficiently large number of such hotels to produce reliable statistics. However, in many cases the user can be provided with data for quite similar hotels that satisfies his needs. In fact, user requests to the AHRP database are frequently very imprecise and sometimes difficult to convert into an exact query.

Arithmetic mean or median data are rarely useful without accompanying information. Generally, one would also like to have an indication of the variability of the sample or population and the number of observations on which the mean was computed. Such information facilitates the identification of significant differences and, other things being equal, helps define the confidence that one can place in the data.

5.4.1 The heuristic model

Traditional static forms of publications, even when they are posted on the Internet, do not allow the flexibility and interactivity which is necessary to build a system which adapts to individually varying information needs. Therefore, the author developed a system which offers this functionality by using a multi-attribute weighting model. This is an heuristic modelling approach as the weighting is basically derived by expert judgements (Wöber, 1999). The objective of this case study was to introduce a system which could identify similar hotel and restaurant businesses and simultaneously guarantee at least a minimum level of representation decided on by the user. This work was also the starting point for research on the optimal selection of comparison partners in business performance studies that led to the present book.

Recalling the extended transformation model from Chapter 3, the main variables identified in this model were the output Y, the discretionary input X^D, and the non-discretionary input X^N. There are several possible ways the relationships between these constructs can be measured and used for the identification of optimal (best practice) comparison partners.

When there seems to be prior knowledge of what is a more favourable environment, caused by X^N combinations, the identification of an optimal comparison partner could be evaluated by finding an optimal $X^D \rightarrow Y$ relationship. The investigation of the environmental factors concerns questions such as which factors are decisive and how many industry sectors (markets) must be distinguished. This evaluation is obviously a stratification problem which needs to consider the (discretionary) input/output relationships. Therefore, all methodologies capable of handling this clustering problem are relevant for the optimal selection of comparison partners.

The stratification of surveys is in part to form groups as homogeneous as possible so that, where numbers allow, analysis of a stratum can occur and otherwise, based on estimates of stratum size, the reliability of the information obtained on aggregation is increased. The AHRP survey uses six criteria (S) in the weighting approach to establish homogeneity while preserving sample size. The stratification is done on the basis of the type of services offered by the establishments, S_1 (Table 5.5), the number of days of operation, S_2, the geographical area where the enterprise is located, S_3 (Table 5.6), the ownership, S_4 (Table 5.7), the size, S_5 (measured by turnover and number of employees) and the category, S_6 (Table 5.8).

Selection of the comparability values is a crucial decision because it has a significant influence on the performance of the decision support system. There is no real theory in hospitality research concerning the criteria that

Table 5.5. Types of service offered by the establishments in the AHRP.

		1	2	3	4	5	6	7
1	Hotel garni	100						
2	Hotel incl. F&B	80	100					
3	Spa hotel	60	80	100				
4	Restaurant	0	50	40	100			
5	Inn, pub, tavern	0	40	30	80	100		
6	Cafe house/shop	0	10	0	50	80	100	
7	Espresso/bar	0	0	0	20	50	80	100

Table 5.6. Geographical area where the enterprise is located.

		1	2	3	4
1	Vienna	100			
2	> 50,000 inhabitants	80	100		
3	15,000–50,000 inhabitants	60	80	100	
4	< 15,000 inhabitants	40	60	80	100

Table 5.7. The ownership.

		1	2
1	Complete ownership	100	
2	Rented or leased	70	100

Table 5.8. Categorization scheme.

		1	2	3	4	5
1	*****	100				
2	****	80	100			
3	***	50	80	100		
4	**	30	50	80	100	
5	*	10	30	50	80	100

determine competitive pressures among hotel and restaurant enterprises. As a result the values had to be assigned by expert judgements in cooperation with the Austrian Professional Hotel and Restaurant Associations and various consultants specializing in hotel and restaurant operations.

A general comparability equation is suggested in the form of an additive function which also includes an additional weighting procedure to adapt to the situational needs of the user accessing the system. The total weighted comparability value C for each individual enterprise i is:

$$C_i = \sum_{j=1}^{6} S_{ij} W_i \tag{5.1}$$

The weights for $w_1, \ldots, w_6$ are assigned by the user according to his requirements by a simple scale from 0 to 3 (0 = not important, 1 = less important, 2 = important, 3 = very important).

After rating each establishment represented in the panel database, the enterprises can be easily sorted by their comparability with the user's case. The number of units under evaluation and the homogeneity in the data set are determined by the number of establishments in the sample, which will be drawn after this sorting procedure. According to the user's desire in the accuracy of the results he will specify a high or low number of establishments which will go into the sample file. This trade-off is made explicit by a bi-polar rating scale where the user is able to decide which of the these conflicting objectives has more importance for the decision problem in hand (Fig. 5.4).

In order to keep the programming effort simple the prototype system uses an integer 9-point-rating ($r = 9$) to decide on the relation between reliability and similarity. The necessary transformation also considers a minimum sample size ($n_{\min}$) of establishments (e.g. 30) to guarantee a certain level of precision. The actual sample size (n) in accordance with the user's importance of reliability (e) is calculated by

$$n = n_{\min} + \frac{e(N - n_{\min})}{r} \tag{5.2}$$

where N denotes the total number of establishments in the panel database.

The respondents to the AHRP study form a sample which is not randomly selected but self-selected. There is little that can be done about this except to make every effort to achieve a high response rate. In Austria some information is available on the structure of the hotel and restaurant industry which is used to weight the sample in order to correct certain sources of bias. For example, in the AHRP study businesses with more employees tend to participate more frequently in the survey. To correct this misdistribution, information about the real size of hotels and restaurants is used; this is available through a regular survey of the Austrian Central Statistical Office.

In a random sample generated from a large population the size of the standard error depends on the size of the sample and is unrelated to the size of the population. For a manager interested in comparing his/her business data with composite figures derived from the database, it is certainly important to have a confidence interval for the estimated industry sector ratio x^* under investigation. For example, when assuming key ratios to be normally distributed, a confidence interval at a 95% significance level ($\alpha = 0.05$) can be defined by

$$\bar{x} - 1.96 \times \sqrt{\frac{\sigma^2}{n}} \le x^* \le \bar{x} + 1.96 \times \sqrt{\frac{\sigma^2}{n}} \tag{5.3}$$

for $\alpha = 0.05$, and $x \sim N\left[0, \sigma_x^2\right]$

The homogeneity of the resulting sample can be expressed by an indicator, derived from Equation 5.1 and expressed in Equation 5.4.

$$C = \frac{\sum_{i=1}^{n} \sum_{j=1}^{6} S_{ij} \times w_j}{6n \times \sum_{j=1}^{6} w_j} \tag{5.4}$$

The indicator, which is standardized between 0 and 100 (100 = complete similarity with the case example entered by the user, 0 = no comparable establishment found in the total panel data set), is displayed together with the key ratios calculated by the program. This homogeneity indicator helps the

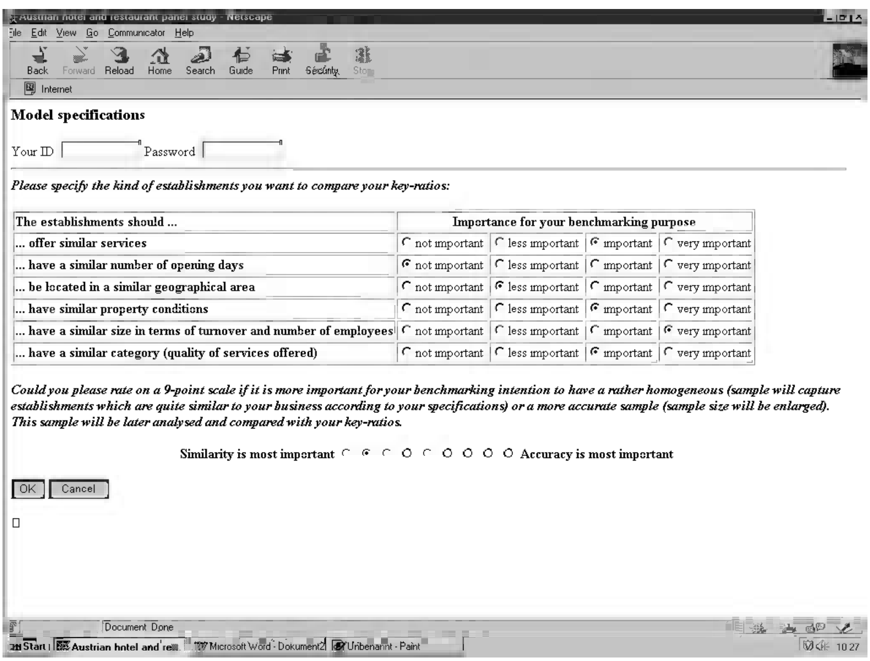
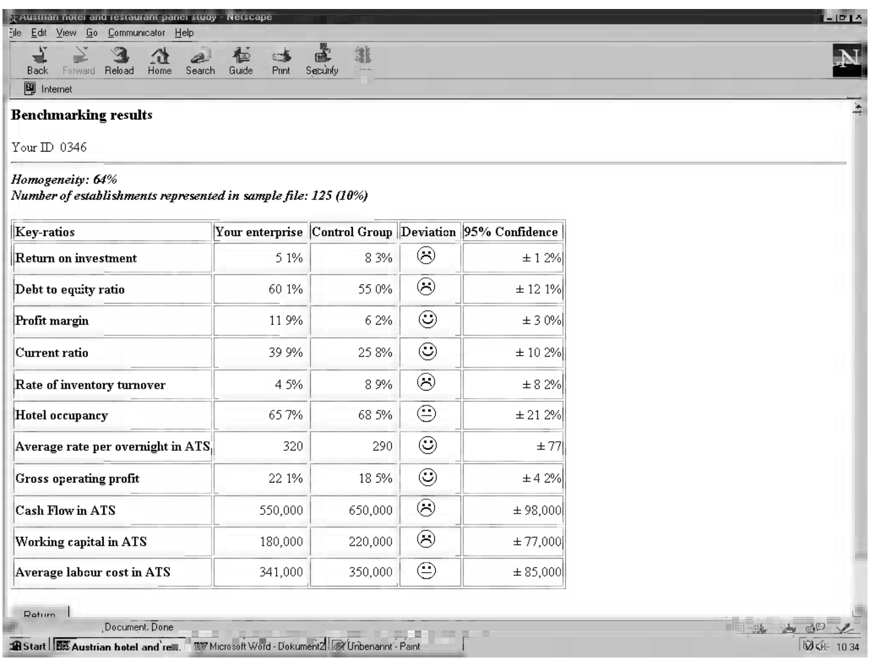


Fig. 5.4. The user decides whether comparability or reliability is more important to his/her decision problem.

manager to understand the composition of the sample and hence supports him/her during the interpretation of the results (Fig. 5.5).

The desired level of precision may be selected by giving the amount of error that someone is willing to tolerate in sample estimates. This amount is determined in light of the uses to which the sample results are to be put. Sometimes it is difficult to decide how much error should be tolerated, particularly when the results have several different uses. In the present application, for instance, an entrepreneur who is interested in opening a new hotel will certainly have completely different information needs in comparison to a manager of an operating establishment. Part of the difficulty is that not enough is known about the consequences of errors of different sample sizes as their effect on the decisions are difficult to observe.

The advantage of the proposed system is that the user can go back and forth and learn from the output. He/she can change the sample size and, therefore, the reliability of the results, as well as the criteria which define the competitive situation he/she is facing. They are not bound to a strict classification as is usual in ordinary printed publications of panel studies. Hence the user will soon realize that results may vary significantly, sometimes even through minor changes in his/her preliminary assumptions. Therefore, he/she can gain more insights and a better understanding of how to interpret benchmarking results and how to use them for managerial purposes.



Benchmarking results

Your ID 0346

Homogeneity: 64%
Number of establishments represented in sample file: 125 (10%)

Key-ratios	Your enterprise	Control Group	Deviation	95% Confidence
Return on investment	5 1%	8 3%	☹	± 1 2%
Debt to equity ratio	60 1%	55 0%	☹	± 12 1%
Profit margin	11 9%	6 2%	☺	± 3 0%
Current ratio	39 9%	25 8%	☺	± 10 2%
Rate of inventory turnover	4 5%	8 9%	☹	± 8 2%
Hotel occupancy	65 7%	68 5%	😐	± 21 2%
Average rate per overnight in ATS	320	290	☺	± 77
Gross operating profit	22 1%	18 5%	☺	± 4 2%
Cash Flow in ATS	550,000	650,000	☹	± 98,000
Working capital in ATS	180,000	220,000	☹	± 77,000
Average labour cost in ATS	341,000	350,000	😐	± 85,000

Fig. 5.5. Sample output page of the hospitality benchmarking program.

There are several caveats to this very simple heuristic approach in the selection of benchmarking partners by the means of financial key ratios, which will be highlighted here and investigated more thoroughly in following chapters. First, a problem relates to the question whether there is the necessary relation between the homogeneity in the company sample and the number of units derived from the panel database. Decisions on the significance levels of confidence intervals and the necessary preciseness in the benchmarking results have a major impact on this problem. Obviously, the adjustments to get reasonably accurate estimates depend on the underlying application which has to be investigated.

Another problem arises during the weighting process of the various competitive criteria. It is clear that the ideal set of weights depends on the decision problem the user is faced with. However, someone could argue that the user might have difficulty in objectively estimating how relevant this is to their decision problem. In fact, first empirical tests of the prototype program showed that users tend to indicate that all criteria are very important for their benchmarking task. Future improvements of the system could incorporate an evaluation of the various decision problems and perform a self-determination of the weighting values. Therefore, poor or irrelevant weights by users could be replaced by system values fed back to users.

Chapter 6

A Mixture Regression Model

The following experiments will compare several alternatives for the optimal selection of benchmarking partners which could probably extend the simple heuristic approach introduced before. The experiments are based on the AHRP database, which will encapsulate both the performance-related issues introduced and the mathematical development necessary to understand the models introduced in Chapter 4. This chapter explores the applicability of the mixture regression model for solving the problem of an optimal selection of comparison partners within the Austrian hotel industry.

One might ask 'Why apply mixture regression modelling for the optimal selection of comparison partners?' First of all, it is obvious that any clustering methodology seems relevant for deriving a subgroup of companies which can be used for performance comparisons. Second, the financial ratios may be appropriate for comparing companies within a sector, but not across sectors, because the ratios for the healthy companies may have different distributions across the sectors. The fortunes of the industrial sectors are bound to be uneven, due to the effects of general market conditions. For example, in some sectors companies tend to have high liabilities because they require substantial capital investment; in other sectors companies require much less extensive capital investment but have higher expenditures that are not reflected in liabilities. It is meaningful to consider 'company' and 'sector' as two sources of variation, and to describe the association of environmental factors and performance by a mixture regression model. In order to uncover these various layers of productivity, conditioned by environmental differences, the mixture regression model seems to be an appropriate instrument.

In consideration of the methodologies discussed in Chapter 4, the following concepts are proposed for the evaluation of comparison partners in benchmarking studies. The approach here will be to allow the logic of the problem to dictate the choice of the methodology instead of first selecting the methodology and then seeking a use for it.

6.1 Sample of the AHRP

For the following tests a subsample was selected which consists only of lodging establishments in the AHRP. Between 1993 and 1999, 6068 accommodation establishments have been compiled in the database,[1] varying between 700 and 900 per year. This represents approximately 3–4% of all accommodation providers in Austria. The decision to exclude restaurants was made because the accommodation and F&B sector distinguish themselves by a variety of characteristics which make a composite treatment within one benchmarking system not impossible but unnecessarily difficult. Secondly, the scaling problems which the author was faced with were particularly significant within the F&B sector. Finally, the decision to solely concentrate on the accommodation sector was further supported by the considerably higher number of lodging establishments in the AHRP.

Unfortunately, not all hotels in the AHRP could be identified by a company number, or similar identification number.[2] Also, hotel chains with consolidated financial statements had to been eliminated for comparability reasons. Hence, one-third of all hotels (33.9%) was excluded at this stage, resulting in a sample of 4013 hotels that could be tracked over multiple years. However, still further accommodation providers' data had to be eliminated from the database. This was primarily due to the high number of missing values in some cases, and/or including companies that did not provide any of the additional information requested on the non-financial questionnaire. In cases of data omissions, the missing data were estimated either from the company's previous years' data (e.g. for missing data on the company's capacity, fire insurance value, balance sheet, etc.) or from figures which obviously could be derived from other data available from that company. For instance, some hotels did not indicate the number of days of operation in their questionnaire. In such cases, sometimes the capacities and the number of generated bednights implied that the hotel was open all year, hence the missing data for the number of days of operation were replaced with 365. All other companies with missing data that could not be estimated as indicated were eliminated from the database.

None of the other procedures commonly used for the replacement of missing values in empirical data sets were applied. This was decided by the author, as some of the following models, especially the non-parametric frontier analysis, are very sensitive to measurement errors which will certainly arise in all available procedures for missing value replacement. In total, 394 (9.8%) companies were excluded due to missing values. The final data set, which included 3619 hotels and similar accommodation establishments, is summarized in Table 6.1.

---

[1] The fiscal years of the data range between 1991 and 1997.

[2] Due to data security reasons, one of the cooperating industry organizations does not provide the information necessary to track individual companies over multiple years.

In Table 6.1 the companies studied are classified into industrial sectors, of which there are 19 in the AHRP. Although the definitions of the sectors were kept stable through the years, companies sometimes changed the main line of their business. The group of spa hotels was eliminated from further analysis due to their unusual size and the importance of side operations (health treatments and medical services).

Table 6.1 shows the asymmetry in the frequency distribution among the various industrial groups. In some sectors there are only a handful of companies, in others there are more than 100. It clearly proves our previous note that larger, high quality accommodation providers (e.g. sectors 2 and 10) tend to participate more in an inter-industry study than smaller providers with lower quality standards (e.g. sectors 7, 9 and 16). Lack of time, knowledge and resources are the main reasons why there is often a resistance by managers of small hotel companies to participate in a benchmarking study (Ogden, 1998).

The bottom of Table 6.1 gives insights to the panel mortality which occurs in the AHRP. The penultimate row ($n^t \rightarrow n^{97}$) notes the number of companies that could be tracked from the respective year until 1997. For example, 187 hotels in the sample can provide business data for 1995, 1996 and 1997, hence their efficiency change could be analysed for 3 consecutive years. The number of companies which can provide data for multiple consecutive years

Table 6.1. Hotel sample used in the study.

Hotel	1991		1992		1993		1994		1995		1996		1997	
sector[a]	no.	%	no.	%	no.	%	no.	%	no.	%	no.	%	no.	%
1	3	0.5	2	0.4	3	0.6	2	0.4	2	0.5	6	0.9	5	1.0
2	128	23.4	120	22.8	127	27.1	142	25.8	107	27.2	159	25.2	144	28.9
3	2	0.4	4	0.8	6	1.3	4	0.7	5	1.3	2	0.3	3	0.6
4	3	0.5	2	0.4	4	0.9	4	0.7	4	1.0	6	0.9	2	0.4
5	82	15.0	84	15.9	75	16.0	93	16.9	55	14.0	93	14.7	76	15.2
6	7	1.3	7	1.3	5	1.1	2	0.4	4	1.0	1	0.2	1	0.2
7	1	0.2			1	0.2								
8	11	2.0	10	1.9	8	1.7	15	2.7	9	2.3	16	2.5	16	3.2
9	2	0.4	1	0.2			1	0.2	1	0.3				
10	113	20.6	112	21.3	70	14.9	111	20.2	84	21.3	136	21.5	92	18.4
11	94	17.2	95	18.0	79	16.8	79	14.4	54	13.7	96	15.2	61	12.2
12	17	3.1	16	3.0	12	2.6	16	2.9	12	3.0	15	2.4	7	1.4
14	18	3.3	16	3.0	22	4.7	22	4.0	19	4.8	22	3.5	23	4.6
15	11	2.0	9	1.7	10	2.1	8	1.5	6	1.5	8	1.3	13	2.6
16	1	0.2	3	0.6	3	0.6	1	0.2	1	0.3	4	0.6	3	0.6
17	37	6.8	30	5.7	30	6.4	28	5.1	21	5.3	37	5.9	23	4.6
18	14	2.6	12	2.3	11	2.3	16	2.9	7	1.8	22	3.5	21	4.2
19	4	0.7	4	0.8	3	0.6	6	1.1	3	0.8	9	1.4	9	1.8
Total	548	100.0	527	100.0	469	100.0	550	100.0	394	100.0	632	100.0	499	100.0
$n^t \rightarrow n^{97}$	61	0.11	75	0.14	95	0.20	132	0.24	187	47.5	352	55.7	499	100.0
$\Sigma n^t \rightarrow n^{97}$	427		450		475		528		561		704		499	

[a]See Chapter 5, p. 88.

decreases with each additional year of observation. Only 61 hotels and similar establishments could provide data for all 7 years of observation. The average panel mortality for the complete time series is 28.8%, varying between 18.7 and 46.9%.[3] The last row of Table 6.1 ($\Sigma n^t \rightarrow n^{97}$) shows the pooled number of companies available for panel analysis, indicating the maximum sample size (704 companies) available for the last 2 years of data.

6.2 Input–Output Variables

For simplification reasons, the experiments to follow are based only on subsets of cases and variables available in the AHRP database. To ensure the comparability of the results among the various methodologies, all analyses were run with companies that have contributed to the AHRP in each of the years between 1991 and 1997. Hence, 61 hotels with operating data for seven years were selected for the analysis (427 cases).

Morey and Dittman's (1995) model was partially adopted to develop the input and output variables for the experiments. Hotel outputs used in the current study are:

1. Y_1: Total F&B revenue (B311 + B312);
2. Y_2: Total accommodation revenue (B314); and
3. Y_3: Average annual bed occupancy adjusted by the number of days of operation (K648).

Multiple output measures have been selected to meet the qualifications of many hotel businesses, i.e. hotel chain companies, where managers of less profitable business units are unhappy at being targeted just on profit, arguing that profitability is largely dictated by factors outside their control, such as the length of time their hotel has been in operation (e.g. hotels in typical summer or winter regions), local population, competition and pitch. Note that the factors suggested by Morey and Dittman and adopted in this study are only examples and could be easily exchanged (or extended) by adding other financial or non-financial factors (e.g. customer response indicators). Also note that factor three was a different scale which will test the universal use of the various methodologies under different scaling conditions.

Factors determined by the market and therefore uncontrollable to the hotel management have been defined:

1. X_1^N: Number of beds (S100);
2. X_2^N: Number of seats in the F&B section of the hotel (S130); and
3. X_3^N: Number of days of operation (S190).

[3] Excluding 1995, which is an unusually small sample caused by problems with one of the cooperating industry organizations, results have an average panel mortality ratio of 24.3%.

The number of days of operation was decided to be uncontrollable as there are many existing hotels in Austria which are annual or seasonal accommodation providers and which were assumed to be constrained by the environment.

Factors controllable by the general manager were defined as:

1. $X_4{}^D$: Total payroll and related (B420);
2. $X_5{}^D$: Material-type expenditures (B410);
3. $X_6{}^D$: Energy costs (B431);
4. $X_7{}^D$: Cleaning costs (B432);
5. $X_8{}^D$: Maintenance costs (B433);
6. $X_9{}^D$: Communication costs (B434);
7. $X_{10}{}^D$: Marketing costs (B435); and
8. $X_{11}{}^D$: Administration costs (B436).

Payroll costs include costs for social insurance contributions, personnel expenses for F&B and housing, seminars, employees' travel expenses, etc. Cost for cleaning captures costs for cleaning materials and rental laundry. Administration costs include office utilities, insurance costs, consultancy costs and travel expenses of the manager. The complete data description is summarized in Tables 6.2–6.8.

There are several limitations that derive from the database. First, it should be noted that the data, values of financial statements and values from an additional questionnaire of all Austrian hotels or other accommodation providers do not involve any sampling scheme or experimental design. These companies are participating in the AHRP on a voluntary basis. Next, there is a

Table 6.2. Business data for 61 Austrian hotels in 1991.

	Mean	Min	Max	SD	%
Number of beds	99	15	394	61.9	
Number of seats	214	24	630	122.5	
Number of opening days	296	157	365	64.4	
Total F&B revenue (€)	726,616	102,178	2,971,229	552,820	
Total room revenue (€)	454,714	30,159	1,883,898	391,287	
Occupancy per opening days (%)	56.3	13.0	94.7	18.6	
Total expenditures (€)	889,276	99,925	3,573,469	696,912	100.0
Payroll and related (€)	419,334	37,499	1,734,483	359,429	47.2
Material-type exp. (€)	232,683	24,927	880,286	162,191	26.2
Energy costs (€)	57,097	7,485	249,413	43,542	6.4
Cleaning costs (€)	21,465	2,471	120,710	21,182	2.4
Maintenance costs (€)	65,745	4,869	305,880	61,853	7.4
Communication costs (€)	16,188	945	77,324	13,736	1.8
Marketing costs (€)	29,294	218	207,554	39,474	3.3
Administration costs (€)	47,470	4,578	229,937	47,188	5.3
Gross profit (€)	292,054	–317,508	1,281,658	240,280	

lack of information about which companies are 'healthy', and which are in bad shape, or are likely to be declared bankrupt in the future, so that we could explore whether the financial key ratios differentiate between such companies. Further, measurement 'error' may be considerable because of uncertainty

Table 6.3. Business data for 61 Austrian hotels in 1992.

	Mean	Min	Max	SD	%
Number of beds	100	15	394	61.8	
Number of seats	208	24	630	111.2	
Number of opening days	297	150	365	63.0	
Total F&B revenue (€)	755,177	133,064	3,116,938	571,165	
Total room revenue (€)	497,567	23,619	2,045,450	421,947	
Occupancy per opening days (%)	58.1	12.2	94.5	19.2	
Total expenditures (€)	955,902	139,168	4,185,156	768,699	100.0
Payroll and related (€)	462,897	53,923	2,020,813	399,401	48.4
Material-type exp. (€)	233,970	24,636	975,415	169,039	24.5
Energy costs (€)	55,377	10,974	214,676	39,028	5.8
Cleaning costs (€)	22,772	3,416	113,806	23,892	2.4
Maintenance costs (€)	80,793	6,977	357,332	80,818	8.5
Communication costs (€)	16,325	945	79,068	13,880	1.7
Marketing costs (€)	31,577	218	231,536	42,844	3.3
Administration costs (€)	52,191	5,741	289,674	52,673	5.5
Gross profit (€)	296,842	–242,509	977,232	203,227	

Table 6.4. Business data for 61 Austrian hotels in 1993.

	Mean	Min	Max	SD	%
Number of beds	104	13	398	63.9	
Number of seats	214	24	630	122.1	
Number of opening days	295	150	365	65.5	
Total F&B revenue (€)	726,915	104,286	3,049,279	560,540	
Total room revenue (€)	503,247	29,214	1,950,176	414,425	
Occupancy per opening days (%)	56.4	13.7	97.0	19.2	
Total expenditures (€)	968,456	142,947	4,056,743	764,916	100.0
Payroll and related (€)	476,903	47,237	2,082,440	400,089	49.2
Material-type exp. (€)	220,812	21,511	901,870	158,863	22.8
Energy costs (€)	57,180	11,410	216,129	39,712	5.9
Cleaning costs (€)	24,465	2,616	157,119	27,397	2.5
Maintenance costs (€)	81,330	5,886	571,935	91,329	8.4
Communication costs (€)	16,494	2,108	69,257	13,267	1.7
Marketing costs (€)	34,217	291	197,815	43,971	3.5
Administration costs (€)	57,054	6,759	283,133	57,936	5.9
Gross profit (€)	261,707	–592,792	942,712	249,452	

about how to assess the business data in the presence of regional differences in inflation, fluctuating prices, various forms of payment for products and services, differences in methods of accounting, and so on. It is obvious, though, that these errors (more appropriately termed 'uncertainties') are likely to be

Table 6.5. Business data for 61 Austrian hotels in 1994.

	Mean	Min	Max	SD	%
Number of beds	105	13	398	62.8	
Number of seats	224	40	630	129.6	
Number of opening days	296	155	365	65.3	
Total F&B revenue (€)	728,905	96,146	2,918,032	547,617	
Total room revenue (€)	517,207	27,979	1,870,526	387,298	
Occupancy per opening days (%)	57.3	28.4	92.9	17.0	
Total expenditures (€)	991,082	144,110	4,039,374	748,574	100.0
Payroll and related (€)	483,635	48,037	1,996,904	384,627	48.8
Material-type exp. (€)	229,180	55,231	937,770	163,151	23.1
Energy costs (€)	59,053	10,683	201,304	42,832	6.0
Cleaning costs (€)	23,538	727	116,858	25,313	2.4
Maintenance costs (€)	81,298	10,102	356,751	79,055	8.2
Communication costs (€)	16,945	2,180	78,414	14,198	1.7
Marketing costs (€)	37,571	145	213,513	49,952	3.8
Administration costs (€)	59,862	6,323	273,105	55,200	6.0
Gross profit (€)	255,029	–766,335	749,184	258,984	

Table 6.6. Business data for 61 Austrian hotels in 1995.

	Mean	Min	Max	SD	%
Number of beds	109	13	398	63.0	
Number of seats	223	25	630	119.1	
Number of opening days	297	120	365	65.8	
Total F&B revenue (€)	742,584	91,858	2,822,250	521,644	
Total room revenue (€)	511,474	37,354	1,892,837	371,649	
Occupancy per opening days (%)	54.9	21.0	90.5	14.1	
Total expenditures (€)	992,688	140,404	4,338,568	759,606	100.0
Payroll and related (€)	487,705	50,217	1,840,003	371,858	49.1
Material-type exp. (€)	210,611	31,540	877,524	148,907	21.2
Energy costs (€)	61,725	10,538	186,333	38,851	6.2
Cleaning costs (€)	23,995	2,471	100,725	22,477	2.4
Maintenance costs (€)	86,894	6,395	681,817	105,582	8.8
Communication costs (€)	16,331	1,962	76,743	13,401	1.6
Marketing costs (€)	39,961	436	304,136	50,698	4.0
Administration costs (€)	65,466	7,413	355,370	69,461	6.6
Gross profit (€)	261,370	–481,094	732,106	194,406	

much larger for companies with large assets and turnover than for small companies. An effective way of bringing these outliers back towards the ‘norm’, is to apply the log-transformation to all variables which might be affected by these scaling-effects.

Table 6.7. Business data for 61 Austrian hotels in 1996.

	Mean	Min	Max	SD	%
Number of beds	110	28	398	61.8	
Number of seats	220	75	630	118.6	
Number of opening days	299	155	365	64.1	
Total F&B revenue (€)	755,610	66,714	2,853,135	534,178	
Total room revenue (€)	509,871	68,094	1,911,078	365,509	
Occupancy per opening days (%)	54.9	18.4	91.4	16.2	
Total expenditures (€)	1,024,046	129,285	4,351,068	764,943	100.0
Payroll and related (€)	506,268	36,482	1,933,533	387,241	49.4
Material-type exp. (€)	214,415	50,726	892,350	151,294	20.9
Energy costs (€)	68,301	10,756	199,632	45,519	6.7
Cleaning costs (€)	25,941	3,416	154,938	29,796	2.5
Maintenance costs (€)	88,406	7,558	630,291	92,246	8.6
Communication costs (€)	14,809	1,526	65,260	12,511	1.4
Marketing costs (€)	38,377	436	252,611	42,176	3.7
Administration costs (€)	67,531	10,756	357,260	73,549	6.6
Gross profit (€)	241,436	–434,438	729,853	221,809	

Table 6.8. Business data for 61 Austrian hotels in 1997.

	Mean	Min	Max	SD	%
Number of beds	109	28	398	61.7	
Number of seats	212	27	630	111.0	
Number of opening days	296	152	365	61.9	
Total F&B revenue (€)	725,719	88,370	2,965,124	516,523	
Total room revenue (€)	518,854	59,592	1,924,013	374,233	
Occupancy per opening days (%)	54.9	19.5	83.5	15.6	
Total expenditures (€)	1,012,975	77,033	4,486,167	768,497	100.0
Payroll and related (€)	504,828	28,851	1,936,949	389,440	49.8
Material-type exp. (€)	212,429	14,825	903,905	152,427	21.0
Energy costs (€)	67,705	5,160	222,960	48,442	6.7
Cleaning costs (€)	23,889	3,270	124,198	25,800	2.4
Maintenance costs (€)	84,244	5,886	686,250	93,577	8.3
Communication costs (€)	12,527	872	61,481	11,352	1.2
Marketing costs (€)	40,137	799	233,934	45,254	4.0
Administration costs (€)	67,215	8,430	401,299	70,266	6.6
Gross profit (€)	231,599	–434,947	805,724	219,173	

6.3 General Procedure for the Selection of Comparison Partners

In mixture regression models it is assumed that a sample on which a measurement is taken is composed of a number of underlying groups or segments. In order to describe the process generating the measurements, a certain statistical distribution is assumed for them. Given one assumed distributional form, the purpose of the mixture approach is to decompose the sample into underlying groups. Whereas in the classical approaches for finite mixtures where only the expected values of each of the underlying densities are estimated, Wedel and DeSarbo (1995) proposed mixture regression models that enable the estimation of the relation of the observations in each underlying group with a set of explanatory variables.

In mixture regression models, as in standard regression models, only one dependent variable can be specified. In order to deal with multiple output problems several mixture regression models have to be defined, analysed independently and combined afterwards.

Applications of mixture regression models to time-series data are relatively rare. Most of the applications provide insights in marketing applications based on demographic datasets, the majority in conjunction with conjoint studies. Response time data were used by Rosenberg *et al.* (1997) who developed a mixture of gamma regressions to describe consumers' gaze duration, registered by eye-movement devices when looking at print advertisements. To the knowledge of the author, so far there have been only a few mixture regression panel data models introduced before.

The general procedure proposed here consists of three major steps:

1. Setting up the specifications of the mixture regression panel data model;
2. Identification of the optimal number of segments; and
3. Evaluation of individual companies.

In the first step the model adaptations allow precautions for the time dependencies which exist in the AHRP database. Statistical tests of the distribution characteristics of the dependent variable are necessary to decide on the model parameters.

In the second step the crucial decision involves establishing the number of segments for the mixture regression model. This should include a discussion on classification problems for non-disjunctive cluster solutions and various test results which compare the mixture regression classification with the traditional AHRP classification system.

In the final step, the evaluation of individual companies takes place. In this core part the regression models identified in the previous steps are used to calculate efficiency scores as described in Chapter 4. Efficient hotels, or 'best practice' companies, are located 'above' the regression line; inefficient hotels below. The distance from a hotel's location to the regression line is calculated and interpreted as an efficiency score. This final step will also include paragraphs emphasizing various aspects of efficiency analysis, i.e. target setting, and illustrate the concept by a case example.

6.3.1 Specifications of the mixture regression panel data model

The program applied for fitting the model was GLIMMIX,[4] a program for estimating mixtures of generalized linear models as outlined before. The program was developed by Peter Boer under the supervision of Michel Wedel (1997, 2001) and was applied in Version 1.0 on a Windows NT computer system. GLIMMIX uses the EM algorithm that sequentially improves upon some sets of starting values of the parameters, and permits simultaneous estimation of all model parameters. The program includes numerous possibilities for model specifications, including type of mixture model, distribution, link function, number of iterations and convergence criterion (Wedel, 2001). Results can be saved into a text file and used for further analysis.

The special specifications that are needed to handle time-series data refer to what Wedel and Kamakura (1999) describe as 'brands' in their consumer research terminology. In their demonstrative examples for GLIMMIX applications, they assume situations where repeated measurements, indexed by $k = 1, \ldots, K$, are taken on $j = 1, \ldots, n$, subjects. Here the study subjects are replaced by companies and repeated measurements are taken for several time periods. In principle, this looks like a relative easy adaptation, however, it needs some further refinement. The problem is that Wedel and Kamakura's applications involved repeated measurements where each subject's k measure was unrelated to any other k measure for the same subject. Their applications did not expect any trend in the sequence of k measures, as must be expected in the AHRP database. With regard to these time dependencies, an additional independent variable X_{12} is introduced to the following models. This variable serves as a moderator for any trends concealed in the database and is defined by the period number, indexed by t.

GLIMMIX requires the dependent variables (Y_1–Y_3) to be normally distributed. The scaling difference between Y_1, Y_2 and Y_3 has consequences on the type of distribution. Revenue figures, being strongly skewed towards the positive, did not fulfil this normal distribution requirement, whereas occupancy rate, being scaled within the range 0 and 1, did (see Table 6.9). However, after a log-transformation of Y_1 and Y_2, both variables passed the Kolmogorov–Smirnov goodness of fit test. Hence, Y_1 and Y_2 were introduced in log-format into the mixture regression model and normal distribution could be selected for the specification of the distribution function with the corresponding identity link function for all three models.

[4] GLIMMIX is a product of ProGAMMA (www.gamma.rug.nl/), a software company closely related to the University of Groningen (The Netherlands).

Throughout all GLIMMIX runs the convergence criterion was set to 0.00001 and the number of EM iterations was chosen to be 150 (the number of minor iterations within the M-step of the algorithm was set to 25).

6.3.2 Identification of the optimal number of segments

For each dependent variable multiple runs were performed in order to find the appropriate number of latent segments in the data. Each model was run with two to four segments, and evaluated by the Bayesian information criterion (BIC).[5] In those results four segments were proposed to be the optimal solution for the given model and another two runs using five and six segments were performed. None of these cases suggested the use of more than five segments, thus further runs with more than six segments were omitted. Table 6.10 gives a summary of these GLIMMIX runs.

The number of segments were defined using the BIC criterion. Therefore, models 1 and 2, using total F&B revenue (Y_1) or total accommodation revenue (Y_2) as an independent variable, reached a minimum BIC value with five latent segments in the GLIMMIX runs, whereas model 3, using occupancy rate (Y_3), achieved its optimal number of segments with three classes.

The occurrence of local optima is a serious problem in the EM algorithm. To investigate the presence of local optima, 15 reruns for each of these optimal

Table 6.9. Log-transformation for Y_1 and Y_2.

	Kolmogorov–Smirnov goodness of fit test[a]				
	Y_1	Y_2	Y_3	$\ln(Y_1)$	$\ln(Y_2)$
1991	0.030	0.069	0.972	0.773	0.927
1992	0.077	0.022	0.967	0.995	0.966
1993	0.071	0.141	0.925	0.913	0.986
1994	0.059	0.082	0.696	0.886	0.987
1995	0.173	0.248	0.983	0.832	0.797
1996	0.087	0.220	0.844	0.713	0.990
1997	0.085	0.253	0.866	0.406	0.510
Mean	0.083	0.148	0.893	0.788	0.880

[a]Tested distribution: normal; values are two-tailed sign.

[5] One of the principal decision-making problems faced by applied statisticians is that of choosing an appropriate model from a number of competing models for a particular data set. The most popular way to solve this problem is to use an information criterion (IC) to make the choice. In general, an IC model-selection procedure is based on choosing the model with the largest maximized log-likelihood function minus a penalty function, which depends on the number of parameters and, in most cases, the sample size. Among a large number of information criteria, Schwartz's (1978) Bayesian information criterion (BIC) and Akaike's (1973) information criterion (AIC) are the most popular.

segment solutions were performed. If different starting values yield to different optima, Wedel suggests selection of the solution with the maximum value of the log-likelihood (Wedel, 1997: 8). The likelihood estimates and the corresponding BIC values for the 15 GLIMMIX reruns for each of the three models are summarized in Table 6.11.

For Y_1, from all 15 runs the eighth run achieved the highest log-likelihood and was therefore selected for further analysis. For Y_2 and Y_3 the best results could be reached in runs 9 and 10, respectively.

Table 6.10. Summary of mixture regression findings with GLIMMIX.

	Segment	AIC	CAIC	MAIC	BIC	R^2
Y_1	2	172	290	201	290	0.84
	3	11	190	55	190	0.89
	4	–66	173	–7	173	0.90
	5	**–275**	**25**	**–201**	**25**	**0.92**
	6	226	–274	–185	87	0.95
Y_2	2	273	391	302	391	0.84
	3	86	264	130	264	0.92
	4	22	261	81	261	0.95
	5	**–87**	**213**	**–13**	**213**	**0.96**
	6	–87	274	2	274	0.96
Y_3	2	–629	–511	–600	–511	0.66
	3	**–720**	**–541**	**–676**	**–541**	**0.76**
	4	–770	–531	–711	–531	0.81
	5	–799	–499	–725	–499	0.85
	6	–852	–490	–763	–491	0.86

Note: Figures in bold signify the segmentation solution derived from the BIC criterion.

Table 6.11. GLIMMIX test (AHRP 1991–1997): 15 reruns.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y_1	LL	175	177	140	191	179	142	196	**208**	165	172	120	174	163	200	170
	BIC	99	94	169	66	91	165	57	**32**	118	105	209	101	122	49	109
Y_2	LL	117	91	78	112	67	125	61	112	**126**	110	105	101	123	91	83
	BIC	215	266	292	225	313	199	326	225	**197**	229	238	246	202	266	283
Y_3	LL	401	390	408	406	387	384	355	391	401	**411**	399	402	408	402	392
	BIC	–536	–513	–549	–546	–508	–502	–443	–516	–535	**–556**	–531	–537	–550	–538	–517

Note: Figures in bold signify the GLIMMIX run recommended by the BIC criterion.

Classification problems for non-disjunctive cluster solutions

Agreement between the resulting classifications will lend more weight to the chosen solution. In order to compare the classification of a mixture regression model with different starting values, someone could compare the posterior probabilities of each of these runs. There are two problems which arise in this context. First, there is the problem of varying labels assigned by the GLIMMIX output. Segment (class) number 1 in iteration 1 is not necessarily class number 1 in the next iteration, which makes comparisons of several GLIMMIX runs (among identical models) difficult.

The second problem is related to the strategy that is applied to the classification problem. This is illustrated by Table 6.12 which shows an ambiguous classification which is valid for all kinds of non-disjunctive cluster solutions. The example shows seven non-disjunctive cluster solutions for one company achieved by the same model and cluster methodology, but employing different starting values. The posterior probabilities are already correctly labelled and the segment assignments, listed in the column 'classification', are therefore not affected by the problem of changing segment labels. In iteration 1 the posterior probabilities suggest classifying the company into segment 1 ($0.8 > 0.2$); in iteration 2 it is classified to segment 2 ($0.4 < 0.6$), and so on.

Nevertheless, the example illustrated in Table 6.12 is still a rather ambiguous clustering solution. The company in question is classified three times into segment 1 (43%) and four times into segment 2 (57%). In order to decide on an exact class someone could consider the number of assignments to each of the two segments. Following this strategy (1) the company is classified to segment 2, which in the present context means that this company will be compared with other companies in this group.

Table 6.12. Example of an ambiguous classification problem for non-disjunctive cluster solutions.

Iteration	Segment 1[a]	Segment 2[a]	Classification	Class 1	Class 2
1	0.80	0.20	1	0.80	
2	0.40	0.60	2		0.60
3	0.70	0.30	1	0.70	
4	0.45	0.55	2		0.55
5	0.90	0.10	1	0.90	
6	0.40	0.60	2		0.60
7	0.40	0.60	2		0.60
				240	235
			Strategy 1:	0.43	0.57
			Strategy 2:	0.80	0.59

[a]Posterior probabilities.

Another strategy (2) considers the values of the posterior probabilities. Using these a company is assigned to the two groups with the highest average posterior probability. The average likelihood of the company being classified in segment 1 is 80%, compared to 59% in segment 2 which clearly favours class 1 over class 2. Although the example in Table 6.12 is a very extreme case, it demonstrates the practical problems with repeated non-disjunctive cluster solutions. On the other hand, agreement between the resulting classifications using both strategies lends support to the case for the validity of a solution.

For the present study the average posterior probabilities for each of the segments identified and both strategies are summarized in Table 6.13. As can be seen from Table 6.13, both strategies lead to the same classification and therefore to the same average posterior probabilities. The high values of the average posterior probabilities indicate that the classification of all companies is very sharply defined.

Mixture regression classification versus AHRP strategic groups

After having achieved several regression models for each latent class in the database, the composition of these groups can be analysed. If the groups are more or less comparable to the traditional AHRP classification system (industry sectors), then the mixture regression model could be replaced by ordinary multiple regression analysis. If the groups are significantly different from the traditional system then this will justify the additional effort.

Each of the attributes which make up the AHRP classification in the hotel group, i.e. location, number of days of operation and category, were tested

Table 6.13. Average posterior probabilities in models Y_1, Y_2 and Y_3.

	Segment	Strategy 1[a]	Strategy 2[a]
Y_1	1	0.9951	0.9951
	2	0.9916	0.9916
	3	0.9792	0.9792
	4	0.9943	0.9943
	5	0.9992	0.9992
Y_2	1	0.9896	0.9896
	2	0.9862	0.9862
	3	0.9916	0.9916
	4	0.9878	0.9878
	5	1.0000	1.0000
Y_3	1	0.9706	0.9706
	2	0.9647	0.9647
	3	0.9983	0.9983

[a]Posterior probabilities.

against the classification derived from the mixture regression models. Only a few relationships could be found. For example, in the occupancy rate model the classes seemed to reflect to some degree differences in the number of days of operation, hence reflecting the industry's adaptability to the seasonal variations in Austria. In this model segment 2 is dominated by companies which are open all year (63%), whereas in groups 1 and 3 the majority of hotels close at least in the pre- or post-season, or open only in summer or winter (76 and 70.7%).

Service quality seems to be an important attribute reflected by the mixture regression classification in model 2 (accommodation revenue group). Here, in segments 1 and 2 the majority of companies are four and five-star category hotels, and thus can be described as the luxury hotel segment. Segment 3 has most of the three-star hotels and segment 4 all the remaining forms of low-budget accommodation.

Although these findings suggest a few similarities between parts of the traditional AHRP classification system and the *a posteriori* defined classes, they are not convincing enough to reject the conditional mixture approach.

6.3.3 Evaluation of individual companies

Following the procedure introduced earlier, the next step involves the evaluation of individual companies. In regression analysis, being a central tendency methodology, the prediction of 'best practices' is performed by projecting companies against the regression line. In the present mixture regression models each of the various segments have a different number of companies and regression coefficients assigned to them. Following the suggestion of Wedel (1997) the selected models are the GLIMMIX solutions with the highest log-likelihood given in Table 6.11.

Interpretation of regression coefficients

The coefficient estimates and the size of each segment for all three models for Y_1, Y_2 and Y_3 are summarized in Appendix Tables A.1, A.2 and A.3.

Overall, only a few coefficient estimates show a high value compared to the standard errors resulting in significant T-values. However, this is not an unexpected result as the selected models are very simple and do not claim to simulate reality. In this respect, simplicity was preferred over authenticity.

What is interesting about the coefficient estimates, however, are the differences among the various classes and models. First, the mixture of regression coefficients for a class in one model cannot be clearly identified in another model. This result is not very surprising as we did not expect mixture regression models to handle multiobjective problems. This observation is also stressed by the fact that model 3 identified three segments, whereas models 1 and 2 detected five. Second, the coefficient estimates vary considerably among the various classes of each of the models, thereby giving each class its individual characteristics. For example, Fig. 6.1 illustrates the regression coefficients for model 3 where occupancy rate was selected as the dependent variable.

Each of the classes in model 3 has a typical coefficient mix with at least one or more coefficients striking out among the values of the other coefficients of the other classes. For instance in class 1, which accounts for 23.6% of all companies in the dataset, administration costs (X_{11}) are more decisive for occupancy rates than in all other groups in model 3. Note that the coefficients are unstandardized, and therefore it does not mean that administration costs are generally more important than other cost factors in relation to occupancy rates. What can be stated is that for companies in segment 3, administration costs have a considerably different influence on occupancy rates than in all other segments, thus justifying the recognition of these companies as a single group.

Efficiency analysis

Following the efficiency approach based on central tendency, each company is projected against the performance standard set by the regression line by putting the company's real data into the regression model. The regression model for $\ln Y_1$, $\ln Y_2$, and Y_3 is defined by:

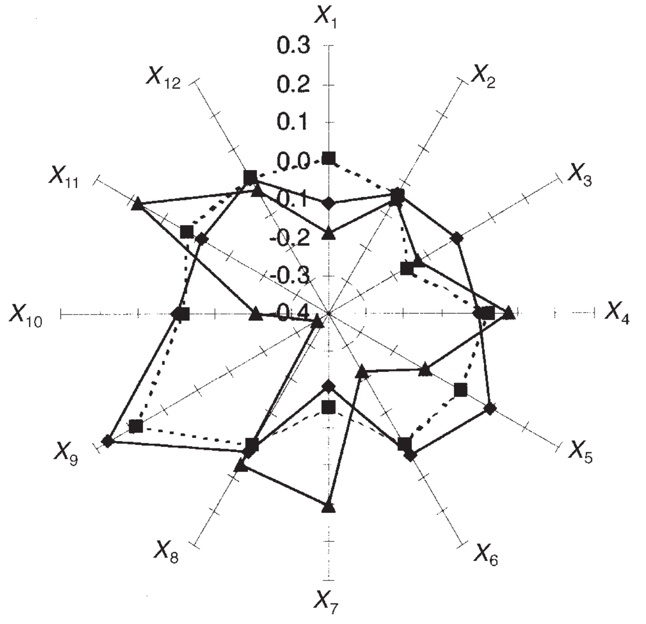


Fig. 6.1. Mixture regression coefficients for model 3 (occupancy rate).
—◆— Class 3 (39.8%); ···■··· Class 2 (36.6%); —▲— Class 1 (23.6%).

$$\hat{y}_{kt} = \sum_{i=1}^{5}\left(a_i + \sum_{j=1}^{11} b_{ij} x_{jkt}\right) p_{ik} \qquad (6.1)$$

where x_{jkt} denotes the jth input variable of the kth company in period t, b_{ij} is the corresponding regression coefficient of segment j, and p_{ik} is the probability that company k belongs to segment i.

Thus, the efficiency scores are calculated by subtracting the estimated from the observed output and can be transformed to a ratio by:

$$e_{kt} = \left(y_{kt} - \hat{y}_{kt}\right) / y_{kt} \qquad (6.2)$$

Efficiency scores are calculated for each individual company, for each year of observation and each individual model resulting in a total of 1281 efficiency scores that are summarized in Appendix Table A.4.

In total, in 636 cases (49.6%), hotels are determined to be efficient, whereas in 645 cases (50.4%) they are inefficient. Within each model, the distribution of efficient and inefficient companies is quite similar. In the F&B revenue model 214 (50.1%) companies are judged efficient, whereas 213 (49.9%) companies are identified as inefficient; in the accommodation revenue model 210 (49.2%) companies are efficient and 217 (50.8%) are inefficient. Finally, in the occupancy ratio model the number of efficient companies is 212 (49.6%), whereas the inefficient companies total 215 (50.4%). What is relatively obvious from this observation is that the regression lines divide the data set into two groups of almost the same size.

Case example

Consider the following representative hotel (no. 2235 in the database) illustrated in Fig. 6.2. Although there was a significant decline in gross profit figures from 1993 to 1995, its level of profit was excellent during almost all years compared to the average gross profit registered by all companies (see Tables 6.2–6.8 and far right column in Table 6.14).

Given the above situation, there are several questions that could be addressed.

- Is hotel no. 2235 operating relatively efficiently and, if not, what is its efficiency rating compared to the other hotels in the database?
- Which of the 60 other hotels would be suitable benchmarking partners for company no. 2235?
- What are concrete targets for each of the eight controllable cost expenditures, if it were as efficient as the average-efficient, among those matching on characteristics derived by the regression models?

Consider the answers to the above questions from the solutions to the mixture regression models as captured in Equations 6.1 and 6.2. In terms of F&B revenues, hotel no. 2235 actually achieved €544,029 in 1997. The mixture regression analysis estimates that, in order for hotel no. 2235 to be rated as efficient in 1997, its total F&B revenues would have been €626,365. The rating of no. 2235 is therefore:

$$\frac{544{,}029-626{,}365}{544{,}029}=-15.1\% \tag{6.3}$$

Company no. 2235, which is a 150-bed hotel, is a member of group four in the F&B revenues mixture regression model. Of all hotels in the database, 19.4% are affiliated with this group with an average efficiency score of −2% in 1997. Note that the average efficiency score in this group is negative, which is a consequence of the statistical variance in the database, rather than by class 4 being a more inefficient class than the other segments. Actually, the central

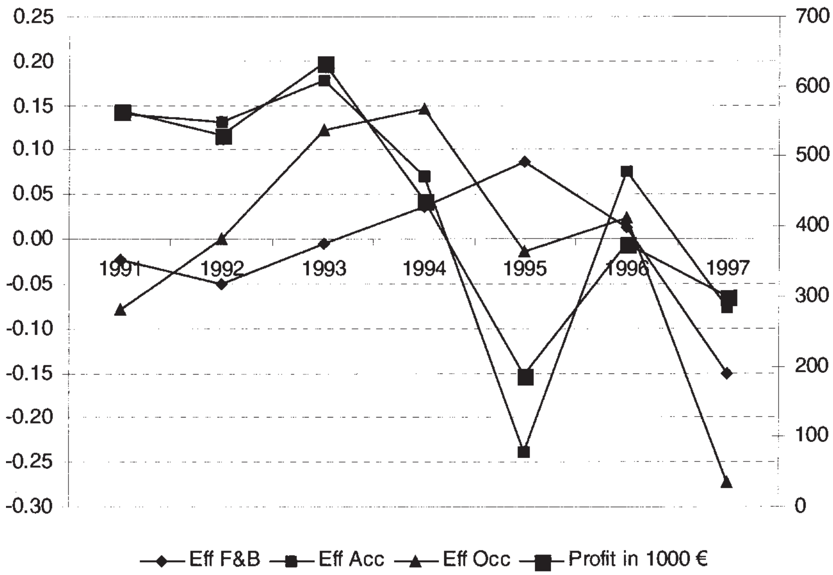


Fig. 6.2. Performance of company no. 2235 between 1991 and 1997.

Table 6.14. Performance of company no. 2235 between 1991 and 1997.

t	F&B revenue	Accommodation revenue	Occupancy	Profit in 1000 € for no. 2235	Profit in 1000 € for 61 Hotels
1991	–0.022	0.141	–0.079	565.8	292.1
1992	–0.051	0.132	0.000	531.3	296.8
1993	–0.004	0.178	0.122	633.7	261.7
1994	0.036	0.070	0.147	435.9	255.0
1995	0.087	–0.239	–0.015	186.8	261.4
1996	0.012	0.077	0.024	374.5	241.4
1997	–0.151	–0.077	–0.273	298.6	231.6

tendency assumptions inherent in the regression methodology affect the results considerably and prevent the formulation of concrete management targets as requested by question three above.

To illustrate this concern for the optimal selection of benchmarking partners, Table 6.15 displays each of the 11 facets of hotel no. 2235 and those of the best practice hotels in the corresponding group 4 of the F&B revenue model 1997.

Among 12 companies in segment 4, hotels no. 2914, no. 1617, no. 2995 and no. 81 which have efficiency scores > 0 were specified as 'best practice companies' and thus were selected as comparison partners. Note that in the proposed solution the selection of these hotels does not depend on company no. 2235, the hotel under review; these four hotels will also serve as benchmarks for other inefficient hotels in segment four.

Table 6.15 makes it clear that using a mixture regression model does not yield clear direction for management action, e.g. target setting. First, it is difficult to decide what and from which of the suggested benchmarking companies hotel no. 2235 should learn. Hotel no. 81 has the highest efficiency rate (12%) and gross profit, hence offering itself as a candidate for comparison. However, by having 33% more capacity in the food section, it seems intuitive that the F&B area has a completely different importance for this hotel. Hotels no. 2995 and no. 2914 are units with less favourable characteristics in terms of restaurant

Table 6.15. Analysis for hotel no. 2235.

Hotel		no. 2235	no. 2914	no. 1617	no. 2995	no. 81	BM[b]
e		–0.151	0.018	0.033	0.118	0.120	
Y_1	€	544,029[a]	698,386	486,036	627,385	1,012,769	706,144
Y_2	€	422,084	684,505	339,746	372,957	727,818	531,257
Y_3	%	49.7	62.9	66.0	34.5	48.5	53.0
X_1		150	92	118	80	140	108
X_2		150	120	170	140	200	158
X_3		160	265	300	365	240	293
X_4	€	215,475	471,937	204,211	493,303	629,856	449,827
X_5	€	127,396	144,692	134,517	161,770	209,152	162,533
X_6	€	47,601	52,979	33,648	55,377	68,240	52,561
X_7	€	16,787	22,965	14,099	7,849	24,636	17,387
X_8	€	95,201	68,749	56,903	51,961	103,486	70,275
X_9	€	12,645	61,481	8,139	11,700	11,846	23,292
X_{10}	€	69,548	5,887	13,881	25,799	62,208	26,943
X_{11}	€	82,847	112,934	39,025	32,703	97,018	70,420
Profit	€	298,613	441,269	321,359	159,880	534,145	364,164

[a]Estimated F&B revenues: €626,365.
[b]Mean values for best-practice hotels in group 4.

capacities compared to company no. 2235, but there are other caveats combined with them. Hotel no. 2995 disqualifies itself by having poor profit figures which are even outperformed by the company under evaluation; hotel no. 2914 has the lowest efficiency score of all four comparative companies.

Obstacles to using mixture models fall into different classes. Benchmarking studies based on mixture regression models depend on the limitations of the standard mixture regression models when distinguishing between controllable and uncontrollable variables, in other words differentiating between constraints and conditional variables. Another obstacle is that serious problems arise when managers operate without guidance in aggregating the 'best practice partners' in order to achieve a set of target values for future hotel management. One option to address aggregation problems is to weight the respective companies with their efficiency values, or to use their final set of posterior probabilities which decided on the group membership. However, both strategies are more intuitive than theory driven, and thus they will not be further investigated.

The need for a multiobjective approach in efficiency and performance studies is well documented by Table 6.16, which shows the results from a correlation analysis between the efficiency scores of the three models and a composite score which was generated by simply summing $e(Y_1)$, $e(Y_2)$ and $e(Y_3)$ for each individual company in the database. The efficiency scores of all three models show a high (and positive), thus plausible, influence on the observed profit of the companies in the database. An indication of the need for multiobjective methods for the selection of comparison partners can be derived from the even higher correlation between the composite efficiency scores, denoted as e(SUM), and the gross profit generated by the companies. The relationship between these composite efficiency scores and the gross profit variable is plotted in Fig. 6.3.

6.4 Limitations of Mixture Models for the Selection of Comparison Partners

Several severe problems in conjunction with mixture models have already been discussed in Chapter 4. These problems arise from limitations associated

Table 6.16. Correlation of efficiency scores from the mixture regression models.

	$e(Y_1)$	$e(Y_2)$	$e(Y_3)$	Profit
$e(Y_2)$	–0.063			
$e(Y_3)$	0.120	0.302[a]		
Profit	0.295[a]	0.192[a]	0.193[a]	
e(SUM)	0.608[a]	0.585[a]	0.728[a]	0.359[a]

[a]Pearson correlation sign > 0.01; n = 427.

with the EM algorithm, such as its sensitivity to local optima and the identification problems due to multi-collinearity in the predictor variables. Other problems include choosing the optimal parameter specifications and specifying the appropriate number of segments.

In the following, specific problems are discussed that only occur in the application of mixture regression models to discover comparison partners in panel databases.

6.4.1 General criticism of the central tendency method

One principal criticism is related to this property of central tendency. Mixture regression analysis involves statistical estimation of the parameters of a fitted line. The criterion for fit is that the line reflects the general structure of the data, so it tends to go through the centre of the data points. When measuring efficiency as a company's distance from the regression line, this 'best practice' line will always result in similar numbers of efficient and inefficient companies. In this respect, using mixture regression models does not differ substantially from using ordinary calculations of mean key ratios as usually done in traditional benchmarking studies. Their advantage, compared to the latter technique, lies in the capability of the analyst to consider scale differences which are determined by the slope of the regression line, hence, yielding individual efficiency scores that are more realistic than those obtained from just subtracting a company's key ratio from one single mean ratio generated by all companies in a specific group.

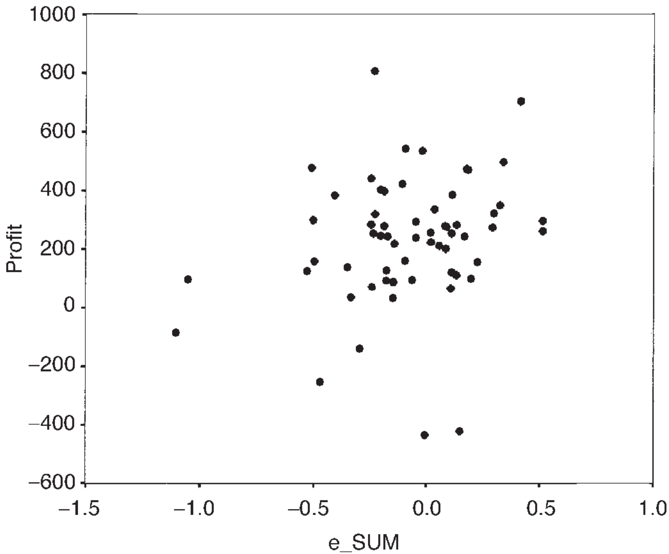


Fig. 6.3. Efficiency scores versus gross profit (composite model).

Under mixture regression analysis the company efficiency is assumed to follow a probability distribution which belongs to the exponential family. The universal property of mixture models simplifies the specification phase prior to analysis and is a real advantage of this technique compared to ordinary regression models.

Successful measurement of efficiency also depends on fitting the right functional form of the curve to the data. Fitting a wrong shape yields incorrect efficiency measures. As in ordinary regression analysis, mixture regression analysis requires prior specification of the shape of the curve to be fitted since it assumes a linear relationship between input and output variables. However, this may be too restrictive and carries the risk of fitting a curve to the data that has the wrong shape. Testing for the appropriateness of the functional form of the curve and choosing alternative models until a better shape is achieved addresses this problem but can result in a computationally very cumbersome procedure.

6.4.2 Disregarding different characteristics of input variables

As mentioned previously, the true efficiency of a company is not directly observable, so it has to be inferred from variables that are observable, such as input and output variables. Assume that the quality of the approximation is sufficient and the analysis shows inefficiencies for a given company. It follows that analysis based on the model should give indications of unprofitable management by comparing the input resources of the given company with the efficient companies in the same mixture regression group. This provides the basic process of target setting. The implication that targets can be set is inherent in comparative studies.

The obstacles that arise in target setting with mixture regression analysis as the underlying methodology arise from severe problems related to distinct treatment of discretionary and non-discretionary variables. In mixture regression models all independent variables are supposed to be discretionary. This limitation leads to situations where objectives set by the best-practice companies obviously cannot be targeted by a specific (inefficient) company due to environmental constraints. This problem is true for both types of non-discretionary variables: for the physical characteristics of a property (e.g. number and mix of rooms) as well as (and especially) for variables determined by a company's market area (e.g. location of a hotel). As proven by the case study of hotel no. 2235 in the present text, the segmentation proposed by the mixture regression analysis did not sufficiently reflect the distinct characteristics of the non-discretionary variables.

However, there is an alternative approach that can distinguish between controllable and uncontrollable variables in mixture regression models. One may define the dependent variable as a key ratio derived from significant output and controllable input factors and include only non-discretionary input factors as dependent variables in the mixture regression model. This model formulation will have the advantage of separating controllable and uncontrollable effects, but at the same time will reduce the benchmarking procedure to a simple clustering of single key ratios. Additionally it will not support the managers' target setting activities, as artificial composite companies made up by best practising companies, which are identified by the mixture regression model, will still vary for a given firm.

6.4.3 Problems with real-time benchmarking systems

In comparison with direct numerical optimization of the likelihood, the convergence of the EM algorithm is slow depending on the data distribution and the initial estimates for the parameters. In the present survey, sometimes more than 150 iterations have been necessary to reach convergence which required between 4 and 6 minutes computation time for each individual run on a Pentium 500 MHz computer system.[6]

This, in combination with other requirements (e.g. restarting the program in order to avoid local optima), makes for a slow analysis process. Applications in real-time systems, like the introduced benchmarking system in Chapter 5, will suffer from severe performance problems. Possibilities to avoid the occurrence of local optima convergence, for instance, the use of simple clustering procedures or *a priori* cluster solutions to obtain an initial partition of the data, may speed up the system significantly.

6.4.4 Inadequacy of multiobjective problems

The analyses of the three mixture regression models have shown that, although the output factors were not contradictory in nature, the results for each of the introduced models presented a very different picture. Although a few hotels had consistently high rankings and a few had consistently low rankings for all three efficiency measures, the majority exhibit considerable variation depending on the indicator chosen. As a result, it is more unclear how to present an overall picture to the manager. The weakness of mixture regression modelling in multiobjective situations is even more clear when someone plans to include more contradictory target variables, for instance, customer satisfaction indicators and expenditures for service operations.

The inadequacy of the mixture regression approach for multiobjective problems is because, using regression analysis, the structure of a system of equations showing the interactions between the many types of output and input variables has to be specified. The calculation of a composite efficiency score by simple summation of all output factors considered, as introduced in

[6] Windows NT (Version 4.0).

the present case study, assumes equal importance of all output factors under evaluation and lacks any substantial theory. A more sophisticated solution could, for instance, foresee a hierarchical system of company objectives. However, despite it sometimes being computationally cumbersome, a complicated system of equations may not be followed easily, making it difficult to apply mixture regression analysis.

Chapter 7

Data Envelopment Analysis

DEA is a mathematical model that was originally used to measure the relative efficiency of operating units with the same goals and objectives. Organizations, similar to operating units, have multiple inputs, such as staff size, salaries, hours of operation and advertising budget, as well as multiple outputs, such as profit, market share and growth rate. In these situations, it is often difficult to determine which organizations/firms are inefficient in converting their multiple inputs into multiple outputs. DEA can separate the efficient firms from the inefficient on the basis of whether they lie on the efficient frontier which is spanned by the best companies in the data set.

In the simple case, DEA analysis is formulated as a fractional programming problem and is then reduced to a linear programming problem as explained in Chapter 4. DEA starts by building a relative ratio consisting of total weighted outputs to total weighted inputs for each company in a given data set. The best organizations in the data set form an 'efficient frontier'. The degree of inefficiencies of other units relative to the efficient frontier is then determined using a linear programming algorithm. An advantage of DEA is that it needs no *a priori* information regarding which inputs and outputs are most important in the evaluation procedure.

The following paragraphs assess the suitability of DEA for the selection of comparison partners in panel databases. In order to achieve results which are comparable with the previous mixture regression, analysis is based on the same database.

The outline of this chapter is as follows. First there is the introduction of a basic input-oriented DEA which will be referred to throughout the chapter. Second, several extensions that have been proposed which address severe shortcomings of the basic DEA model are discussed in detail. One major problem area in the context of the mixture regression approach is the distinct treatment of discretionary and non-discretionary variables. For simplicity reasons, the basic model starts by defining all input factors as discretionary.

However, later on, this issue is subject to a special section on uncontrollable input factors and the possibilities of their integration in DEA. Another extension discussed in the text focuses on the DEA problem with time series data. At the beginning the DEA basic model is not calculated for all data at once but requires separated analysis for each individual year. This does not allow the incorporation of dependencies inherent with time series data. Later on, the two most important methodologies for longitudinal DEA analysis, the 'window approach' and the Malmquist DEA, are both applied to the AHRP database. Finally, the chapter focuses on another modification of DEA which allows the calculation of efficiency scores even for efficient companies, hence, building a procedure for ranking efficient units in the database. As in Chapter 6, the section ends with a summary and discussion of the utility of DEA for the optimal selection of comparison partners.

7.1 Model Specifications

The experiments presented here are based on the same subset of variables, which was introduced in the previous chapter. A total of 61 hotels with similar business characteristics was selected from the AHRP database for the DEA. The operating data entered in the analysis are based on financial information from 1991 to 1997.

The model selected here is an adaptation of the one introduced by Morey and Dittman (1995). It is an input-oriented model, which seeks to identify technical inefficiency as a proportional reduction in input usage. The input-oriented model will use a variable returns to scale (VRS) model as it cannot be assumed that all hotels in the database are operating at an optimal scale. Imperfect competition or constraints on finance (e.g. regional differences in the tax system) are very likely to occur in the Austrian accommodation industry, which may cause a hotel to not be operating at optimal scale. The use of the VRS specification (see p. 74) permits the calculation of technical efficiencies resolved by any scale efficiencies effects.

The basic DEA problem described here involves the three output variables and the 11 input variables introduced. For the moment, the differences between discretionary and non-discretionary variables are neglected. They are, however, in one of the DEA extensions discussed at a later stage of this study. In DEA it is important to keep the number of input variables as small as possible to avoid an overestimation of efficient companies in the database (Li and Reeves, 1999: 507). For this reason the various types of expenditures, X_4 to X_{11}, are summarized to a new discretionary input variable, referred to as 'total expenditures', and denoted with X_4 in the forthcoming models. Finally it is important to note that DEA does not require any distributional assumptions, and therefore the data are entered in their original scale.

The complete DEA model is defined in terms of N linear programming problems to be solved, each for a base organization, resulting in N different weight sets. In each linear program, the number of constraints is held constant and only the ratio to be maximized is changed.

Several special computer programs are used to solve this bulk of linear programming run in one step. Commercial software packages are available from Banxia Frontier Analyst,[1] OnFront,[2] Warwick-DEA,[3] IDEAS,[4] and DEA-Solver.[5] For the General Algebraic Modeling System (GAMS), a high-level modelling system for mathematical programming problems, there is a free DEA code at PARN,[6] described by Olesen and Petersen (1996). Free DEA software products, which can be downloaded from the Web for academic use, are DEAP[7] developed by Tim Coelli (University of New England) and EMS[8] developed by Holger Scheel (University of Dortmund).

This study involves the use of the latter two computer software tools. First, for the basic and time-dependent analysis with solely discretionary input variables, it will employ DEAP (Rel. 2.1) described by Coelli (1996). Later, when the inclusion of non-discretionary variables is discussed, it will employ Holger Scheel's EMS (Rel. 1.3) software described by Scheel (2000).

7.1.1 Analysis of individual years

Originally DEA models were not designed to handle time series data or to evaluate performance from a dynamic point of view. Therefore, the basic DEA model introduced here is based on individual runs for each of the 7 years under evaluation. Recall the input-oriented DEA model with VRS specification by Banker (1984):

$$\min f_o \qquad (7.1)$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{im} \le f_o x_{io} \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jm} \ge y_{jo} \qquad j = 1, \ldots, s$$

$$\sum_{m=1}^{n} \lambda_{om} = 1$$

1 www.banxia.co.uk
2 www.emq.se/onfront1.htm
3 www.warwick.ac.uk/~bsrlu/dea/deas/deas1.htm
4 www.ideas2000.com
5 www.saitech-inc.com
6 www.parn.org.uk/
7 www.une.edu.au/econometrics/deap.htm
8 www.wiso.uni-dortmund.de/lsfg/or/scheel/ems/

In the present application to the AHRP database, there are four input variables i ($r = 4$), and three output variables j ($s = 3$). For each of the 61 companies in the database a linear programming task had to be defined, which made 427 optimization runs in total.

The piecewise linear form of the non-parametric frontier in DEA can cause difficulties in efficiency measurement of some companies which define the frontier. Refer to Fig. 4.8 (p. 70) where the companies using input combinations P_3, P_4, P_5 and P_6 are the four efficient companies that define the frontier, and companies P_1 and P_2 are inefficient firms. However, it is questionable as to whether the point P_6 is an efficient point since one could reduce the amount of input x_2 (until $P_6 = P_3$) and still produce the same output. This phenomenon is known as input slack and can be found in output-oriented models as well (output slack). The DEAP software gives the user three choices regarding the treatment of slacks (Coelli, 1996: 14):

1. One-stage DEA, which solves Equation 4.28 (p. 68), but ignores the need for a second optimization step. Slacks are calculated residually;
2. Two-stage DEA, which introduces the non-Archimedean infinitesimal (see p. 71), a very small number in Equation 4.31 used by most DEA software packages; and
3. Multistage DEA, which conducts a sequence of radial linear programs to identify the efficient projected point (Coelli, 1997).

The advantage of the multistage specification for slack optimization is that it identifies projected efficient points which have input and output mixes that are as similar as possible to those of the efficient points, and it is also invariant to units of measurement (Coelli, 1996: 15). Therefore, this specification was selected for the analysis.

A complete summary of the findings is presented in Appendix Table A.5, which contains the efficiency scores for all seven years and the number of times each company resides on the frontier, and thus was selected as a peer for the inefficient companies. The far right column sums up the peer counts for each hotel and gives an overall picture of an individual hotel's performance. For example, hotels no. 2631 and no. 2788 are the two most preferred benchmarking partners with 90 and 88 times being a peer candidate, respectively. Between 1991 and 1997, among all hotels, a hotel could be identified being inefficient 219 times, which is 51.3% of all observations in the database. The mean score of all inefficient hotels between 1991 and 1997 is 0.829 and varies between 0.467 (no. 396 in 1997) to 0.998 (no. 2287 in 1994).

Figure 7.1 shows the operating profit plotted against the DEA efficiency score for each of the 61 hotels in 1997. The hotels deemed efficient and given a score of 1 are plotted along the right border. Although the relationship between profitability and efficiency is obvious, the plot clearly shows that not all hotels that have been classified as efficient are necessarily also profitable. There are some low-profitability hotels that are run efficiently, and some high-profitability hotels that are run inefficiently. The DEA solution plotted in Fig. 7.1 illustrates that by the use of this methodology, unanticipated insights may be obtained and may thus redirect managerial action. In a broader sense, the DEA framework can create an approach for learning from outliers and for inducing new theories of best practice.

Next, the capabilities and limitations of this simple DEA approach will be demonstrated by a case example.

Target setting

Hotel no. 283 (the eighth hotel in the list of hotels in this study) is a property having 97 beds and was selected for more detailed analysis. The possibilities for using DEA findings for managerial target setting can be demonstrated by means of the 1996 data. In this year, hotel no. 283 achieved €2,081,000 in F&B revenue and €1,061,400 in room revenue, and spent €3,074,600 for operating expenses. The gross profit was €67,700, or 2.2% of the total revenue. The average occupancy rate per number of opening days was 68.9%. That compares favourably with the average set's occupancy rate of 54.9%, F&B revenue of €755,610, and total room revenue of €509,871 in 1996. Hence hotel no. 283 exceeded the average output measures of all hotels in the data set on all three of the performance measures. Variable returns to scale DEA shows that four of the remaining 60 hotels constitute a peer group for company no. 283. However, as is typically the case, the characteristics of the peer-group members do not perfectly match those of the unit being evaluated.

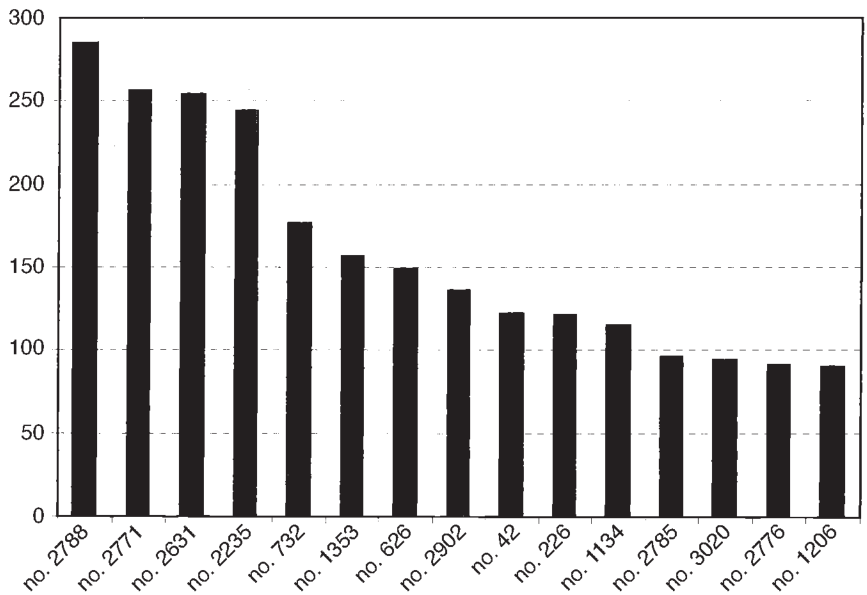


Fig. 7.1. Profitability versus DEA efficiency for 61 Austrian hotels in 1997.

The linear program builds a weighted composite of the four efficient peer-group members identified so that they perfectly match the levels of outputs of the hotel being evaluated. Thus a composite, efficient benchmark company is derived that can be used to develop the targets for hotel no. 283. Particularly, it shows the resource–expenditure targets for hotel no. 283 based on the management achieving at least the same room revenue, F&B revenue, and occupancy rate as benchmark. The benchmarking figures are summarized in Table 7.1.

The efficiency score for hotel no. 283 is 0.948. As it turns out, hotel no. 283 ranks only 40th out of the 61 hotels in efficiency, notwithstanding the fact that the hotel's occupancy rate and revenue figures all exceed those of the average set. The peer group is matched to hotel no. 283 on the levels of all outputs. In relation to output factors the virtual benchmarking company either performs equal or even better than the hotel under evaluation. The benchmarked F&B revenue figures are the same as those for hotel no. 283, and the benchmark for the room revenue as well as the occupancy rate show higher levels compared to hotel no. 283.

The benchmarks that are most interesting are those for the eight resource expenditures. They constitute a set of guidelines for hotel no. 283 to work toward. If hotel no. 283 utilizes its resources more efficiently, it should be able to achieve the same output with lower expenditures. The DEA analysis indicates that hotel no. 283 has the potential to improve its gross profit by

Table 7.1. DEA benchmarks for hotel no. 283.

t = 1996	m = 8 (no. 283)	m = 4 (no. 42)	m = 14 (no. 626)	m = 49 (no. 2788)	m = 43 (no. 588)	BM[b]
Total F&B revenue (€)	2,081,000	1,947,800	2,853,100	484,700	1,348,700	2,081,000
Total room revenue (€)	1,061,400	525,400	1,911,100	68,100	209,500	1,195,700
Occupancy[a] (%)	68.9	80.4	83.0	37.5	42.4	70.5
Number of beds	140	137	180	28	68	133
Number of seats	550	220	400	150	180	304
Number of opening days	292	260	270	300	364	277
Total expenditures (€)	3,074,600	2,097,600	4,351,100	310,500	1,330,800	2,915,300
Payroll and related (€)	1,580,100	877,300	1,933,500	171,200	666,000	1,295,800
Material-type exp. (€)	552,800	580,600	892,400	52,600	378,600	620,200
Energy costs (€)	199,600	108,800	167,700	21,400	83,700	119,600
Cleaning costs (€)	89,200	91,300	79,100	11,000	19,600	63,000
Maintenance costs (€)	210,800	160,700	630,300	17,800	71,700	391,100
Communication costs (€)	50,100	18,400	65,300	6,400	19,900	42,100
Marketing costs (€)	72,200	138,900	252,600	7,300	31,000	168,900
Administration costs (€)	319,800	121,700	330,200	22,800	60,200	214,700
Gross profit (€)	67,700	375,600	413,100	242,300	227,400	361,300

[a]Per opening days; [b]benchmark (composite group) based on LP solution with λ_{42}* = 0.162, λ_{14}* = 0.571, λ_{49}* = 0.259 and λ_{43}* = 0.008; e_8* = 0.948.

€293,600 (more than four times the observed gross profit), when it reduces it expenditures to €2,915,300 (a decrease of 5.2%). The identification of the four peer properties can be made available to hotel no. 283 so that the management can ascertain from them (perhaps through site visits) the details of the processes and practices that enable them to perform better.

Finally it should be mentioned that the target-setting capabilities of DEA demonstrated by the example can be used to develop policy-making scenarios that would enable managers to identify the operating response to different managerial priorities. This kind of sensitivity analysis in conjunction with scenario planning for electricity generating plants was recently introduced by Athanassopoulos *et al*. (1999).

7.1.2 Longitudinal analysis in DEA

Originally, when Charnes *et al*. (1978) developed the basic DEA methodology they did not think about time series applications. This heightens the problem of noise and makes one-off performance assessment highly tentative. The repeated application of DEA through a panel data set, as described in the previous paragraph, produces little more than a continuum of 'static' results. In reality the behaviour underlying the production process is likely to be dynamic because hotel management may take more than one time period to adjust their input factors to the desired output levels. Furthermore, capital inputs have a multi-period dimension since they generate outputs for future periods.

In principle, dynamic aspects of DEA have not been investigated very thoroughly in the past. Yet models in the applied literature are based almost exclusively on current inputs, biasing efficiency comparisons against capital intensive processes. Only recently have a few authors investigated this subject (Charnes *et al*., 1985a, 1994a; Sengupta, 1992, 1995, 1999; Färe and Grosskopf, 1996; Murthi *et al*., 1996; Coelli *et al*., 1998). The two main approaches suggested so far to address this matter are the window analysis technique and the Malmquist DEA. These two methodologies are discussed in the following two paragraphs.

The window approach

Dynamic DEA modelling has to deal with the problems of the trends in data in, for example, growing organizations and inflationary environments. The first initiative to consider DEA for multiperiod analysis was put forward by Charnes *et al*. (1985a) with their application to perform a study on aircraft maintenance operations. They proposed performing DEA over time using a moving average similar procedure, where a company in each different period is treated as if it were a 'different' company. Specifically, a company's performance in a particular period is contrasted with its performance in other periods in addition to the performance of the other companies. So far, the most comprehensive application of the window technique to DEA was a study of brand efficiency among various segments in the US carbonated beverage industry by Charnes *et al.* (1994a).

DEA window analysis considers the same companies in different time periods as separate observations. Out of a total of T years, DEA is performed on all companies defined in the earliest $t < T$ contiguous years, that is, years 1–t. It is then performed on years 2–$t + 1$, 3–$t + 2$, and so on. Each company is thus analysed several times with slightly different comparison sets. In general, Charnes *et al.* (1994a), who applied the windows technique on a quarterly data set, found that $t = 3$ or 4 tended to yield the best balance of quality of information and stability of the efficiency scores.

The windows analyses using the AHRP database were performed for 3- and 4-year windows. The results were essentially identical, which indicates that the analysis of data is robust and not sensitive to the choice of the number of years in the moving window. The results presented here focus on the three-years solution, for which there are five window runs.[9] Table 7.2 shows the results for five hotels taken from the much larger window DEA. The four columns added on the right of Table 7.2 provide diagnostics for the stability of each brand's efficiency ratings. The first two of these columns contain the mean efficiency rating and its variance over the 15 evaluations of each hotel. The third column shows the largest difference in efficiency scores recorded for a single period, and the fourth column gives the difference between the maximum and minimum scores over all evaluations. These results show that the efficiency scores are fairly stable. This was usual for all window analysis runs. A summary of the complete analysis for all hotels is provided in Appendix Table A.6.

ANALYSIS OF PEER APPEARANCE When a DEA is performed for a single period, one might suspect that some of the peer hotels appear there only by chance and that they will not necessarily be a peer on other time periods. With a window analysis, each hotel is evaluated $m \times k$ times, where m is the number of windows and k is the number of periods in a window. Thus there are $m \times k$ lists of peer hotels (with their corresponding optimal λ values) for each hotel in the database.

Charnes *et al.* (1994a) proposed to analyse these results either by counting the number of times efficient companies appear in the reference set of all companies (including its own) or by summing the optimal λ values corresponding to each efficient company over all peers. In order to derive the efficient reference companies that are most important in determining a particular company's efficiency, they established a 'facet participation table' (Charnes *et al.*, 1994a: 157), which allows one to rank order the efficient brands by their overall influence on the reference set. The table they introduced also

9 In general, the number of window runs can be calculated by the number of available periods – the number of periods in the moving window + 1.

gives insights into the geometric properties of the empirical production function. Companies with high counts tend to be located near the centre of the production frontier whereas those with low counts are located near the edges.

The ordering of some hotels starting from the highest peer appearance is displayed in Fig. 7.2. Hotel no. 2788, for instance, is the most frequent reference hotel. This is based on 285 peer appearances. Note that the maximum number of times a hotel may be counted as a peer member is 915 (= 15×61).

Such a graph gives one a perspective on the most robustly efficient companies, i.e. companies that appear as reference firms most often. Based on the pattern seen one can establish a cut-off point in terms of the number of peer appearances and then select those companies that meet or exceed this standard as 'exemplary cases'.

A second possible way to gain insight from the peer statistics, which was suggested by Charnes *et al*. (1994a) but not examined in detail, was to consider the weights of the peers in the analysis. From the perspective of what makes a company inefficient, for instance, greater attention might be paid to those

Table 7.2. Summary of window DEA results for five hotels (t = 3).

j	Hotel no.	1991	1992	1993	1994	1995	1996	1997	Mean	Var	Column range	Total range
1	14	0.835	0.929	0.961					0.969	0.002	0.126	0.165
			0.940	1.000	0.987							
				0.984	0.970	0.995						
					0.983	0.994	1.000					
						0.991	1.000	0.968				
2	25	1.000	1.000	1.000					0.992	0.000	0.049	0.049
			1.000	1.000	1.000							
				1.000	1.000	0.951						
					1.000	0.951	1.000					
						0.983	1.000	0.997				
3	38	0.932	1.000	1.000					0.936	0.006	0.168	0.169
			1.000	1.000	1.000							
				1.000	1.000	0.884						
					1.000	0.865	0.832					
						0.861	0.831	0.832				
4	42	1.000	1.000	0.977					0.979	0.001	0.100	0.100
			1.000	0.978	1.000							
				0.999	1.000	0.903						
					1.000	0.900	1.000					
						0.924	1.000	1.000				
5	81	0.880	0.906	0.857					0.870	0.002	0.099	0.139
			0.910	0.858	0.811							
				0.855	0.813	0.803						
					0.876	0.861	0.890					
						0.876	0.914	0.942				

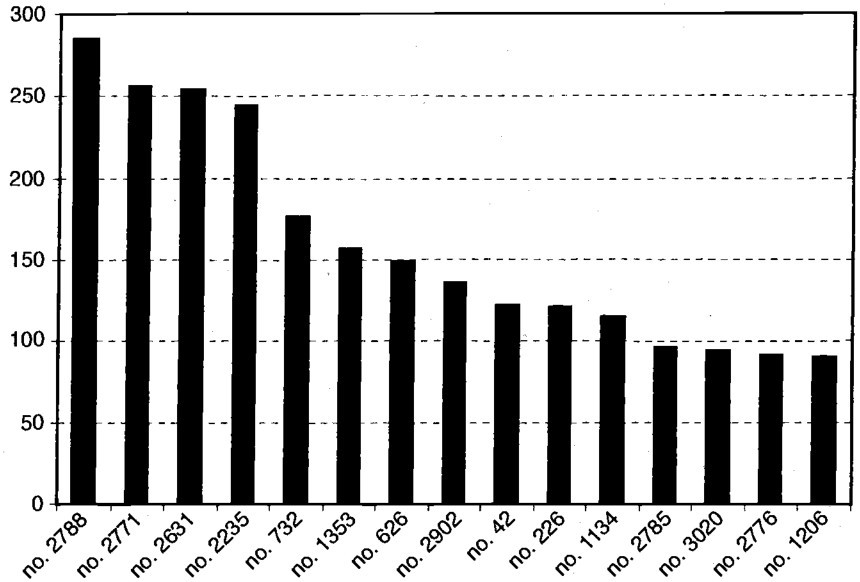


Fig. 7.2. Hotels with the most peer appearances.

peers with the highest λ values in a particular evaluation. In multiple evaluation situations, such as in window analysis, peer members with the largest sums of λ values across the total of their evaluations may be considered for comparative purposes.

For hotel no. 472 this is illustrated in Table 7.3. Between 1991 and 1997, based on the windows analysis, hotel no. 472 was identified ten times as inefficient and five times as efficient (column N in Appendix Table A.7). During the inefficient occurrences the hotels listed in Table 7.3 presented themselves as comparison partners of hotel no. 472. Note that hotel no. 472 itself appeared six times in its own peer group, since, by definition, window analysis allows a company to be compared with its past and future performance. Also note that the mean weights (second column) clearly deviate from the number of counts (third column), thus leading to a different ranking for the optimal selection of comparison partners. Recommendations concerning the fine tuning of partner choice therefore differ depending on the strategy selected.

In summary, efficient companies with the largest sums of λ values across all observations most strongly evaluate all other companies. Because of this they may be considered as the benchmarks industrywide. To allow insights based on this a summary of mean λ values for the present study is included in Appendix Table A.7.

As can be seen from the results presented the window analysis technique is an area for further research extending DEA. For example, the choice of the number of time periods in a window for DEA window analysis is entirely *ad hoc*. The problem of choosing the width for a window (and the sensitivity of

Table 7.3. Peers for hotel no. 472 in 1991–1997.

Peer no.	Mean	n	SD	Min	Max
472	0.583	6	0.388	0.035	0.920
1061	0.472	2	0.031	0.450	0.494
1206	0.302	5	0.194	0.089	0.486
626	0.196	10	0.140	0.016	0.362
2005	0.184	2	0.210	0.035	0.332
1134	0.093	8	0.064	0.008	0.184
921	0.084	2	0.021	0.069	0.098
226	0.079	3	0.053	0.019	0.119
2776	0.072	7	0.058	0.007	0.169
42	0.037	1		0.037	0.037
2631	0.022	1		0.022	0.022
2914	0.007	1		0.007	0.007
2771	0.005	1		0.005	0.005

DEA solutions to window width) is currently determined by trial and error. Furthermore, representing each company as if it were a different company for each period in the window must be replaced by an approach that recognizes the continuity of firms over time. From an overall point of view, the high dependence on intuition involved in using the window analysis is motivation to think about alternative ways of time series analysis with DEA such as the Malmquist DEA.

The Malmquist DEA approach

Färe *et al.* (1992) combined Farrell's ideas of efficiency with some work of Caves *et al.* (1982) on the measurement of productivity to a Malmquist index of productivity change. Caves *et al.* defined their Malmquist productivity index as the ratio of two input distance functions, while assuming no technical inefficiency in the sense of Farrell (1957). Färe *et al.* extended this approach by dropping the assumption of no technical inefficiency and developed a Malmquist index of productivity that can be decomposed into indices describing changes in technology and efficiency. Finally, Färe and Grosskopf (1996) summarized their findings on intertemporal production frontiers in a comprehensive textbook.

Recently, there were two extensions to the original Färe *et al.* (1992) approach. First, Simar and Wilson (1999) give a statistical interpretation to the Malmquist productivity index and its components, and present a bootstrap algorithm which may be used to estimate confidence intervals for the indices. Second, Löthgren and Tambour (1999) modify and apply the DEA concept to model both production and consumption activities in Swedish pharmacies.

Färe *et al.* (1992) specify an input-based 'Malmquist productivity change index' as:

$$M_i\left(M_{t+1},x_{t+1},y_t,x_t\right)=\frac{d_i^{t+1}\left(x_{t+1},y_{t+1}\right)}{d_i^t\left(x_t,y_t\right)}\left[\frac{d_i^t\left(x_{t+1},y_{t+1}\right)}{d_i^{t+1}\left(x_{t+1},y_{t+1}\right)}\times\frac{d_i^t\left(x_t,y_t\right)}{d_i^{t+1}\left(x_t,y_t\right)}\right]^{1/2} \quad (7.1)$$

The notation $d_i^{t+1}\left(y_t,x_t\right)$ represents the distance from the observation in period t to period $t + 1$. Hence, the input-oriented productivity measure compares the input requirements for producing output level y_t, produced in period t, with the input that would have been required if the production technology was the same as in period $t + 1$ (Grosskopf, 1993: 183). This means the input oriented index essentially compares x_t with what would have been required in period $t + 1$. The subscript i in Equation 7.1 indicates the input-orientation of the measures.[10]

Values $M_i(t_1,t_2) < 1$ indicate improvements in productivity between t_1 and t_2, whereas values $M_i(t_1,t_2) > 1$ indicate decreases in productivity from time t_1 to t_2; $M_i(t_1,t_2) = 1$ would indicate no change in productivity. The ratio outside the square brackets of Equation 7.1 measures the change in Farrell's input technical efficiency between periods t_1 and t_2, and defines the input-based index 'efficiency change':

$$e_i\left(y_{t+1},x_{t+1},y_t,x_t\right)=\frac{d_i^{t+1}\left(x_{t+1},y_{t+1}\right)}{d_i^t\left(x_t,y_t\right)} \quad (7.2)$$

Values of $e_i(t_1,t_2)$ less than 1 indicate improvements in efficiency between t_1 and t_2, and vice versa. Similarly, the remaining part of the right-hand side of Equation 7.1 defines an input-based measure of 'technical change':

$$T_i\left(y_{t+1},x_{t+1},y_t,x_t\right)=\left[\frac{d_i^t\left(x_{t+1},y_{t+1}\right)}{d_i^{t+1}\left(x_{t+1},y_{t+1}\right)}\times\frac{d_i^t\left(x_t,y_t\right)}{d_i^{t+1}\left(x_t,y_t\right)}\right]^{1/2} \quad (7.3)$$

As with $M_i(t_1,t_2)$ and $e_i(t_1,t_2)$, values of $T_i(t_1,t_2)$ less than 1 indicate technical growth between times t_1 and t_2, and vice versa.

DEA can be used to measure the distance functions which make up the Malmquist index (Färe *et al.*, 1994). In a two-period case, four distance functions must be calculated for each company in the database, thus requiring one to solve four linear programming problems. Recall the basic input-oriented DEA model for constant returns to scale in Equation 4.29 (p. 69). Including the time subscript gives the first linear program:

$$\left[d_{io}^t\left(x_{io},y_{to}\right)\right]^{-1}=\min f_o \quad (7.4)$$

subject to:

$$\sum_{m=1}^{n}\lambda_{om}x_{itm}-f_o x_{ito}+s_1^-=0 \qquad i=1,\ldots,r$$

[10] Note that the output-oriented Malmquist index is defined similarly to the input-oriented measures presented here (see Coelli *et al.*, 1998).

$$\sum_{m=1}^{n} \lambda_{om} y_{jtm} - s_j^+ = y_{jto} \qquad j = 1, \ldots, s$$

The other three linear programs are logical derivations from Equation 7.4:

$$\left[d_{io}^{t+1}(x_{t+1o}, y_{t+1o})\right]^{-1} = \min f_o \tag{7.5}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{it+1m} - f_o x_{it+1o} + s_1^- = 0 \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jt+1m} - s_j^+ = y_{jt+1o} \qquad j = 1, \ldots, s$$

$$\left[d_{io}^{t}(x_{t+1o}, y_{t+1o})\right]^{-1} = \min f_o \tag{7.6}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{itm} - f_o x_{it+1o} + s_1^- = 0 \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jtm} - s_j^+ = y_{jt+1o} \qquad j = 1, \ldots, s$$

$$\left[d_{io}^{t+1}(x_{to}, y_{to})\right]^{-1} = \min f_o \tag{7.7}$$

subject to:

$$\sum_{m=1}^{n} \lambda_{om} x_{it+1m} - f_o x_{ito} + s_1^- = 0 \qquad i = 1, \ldots, r$$

$$\sum_{m=1}^{n} \lambda_{om} y_{jt+1m} - s_j^+ = y_{jto} \qquad j = 1, \ldots, s$$

For an output-oriented formulation of the linear programs see Coelli (1996: 28). The approach can also be extended to determine a DEA solution based on variable returns to scale. This requires the solution of two additional linear programming problems with convexity restrictions added to each. Compared to the constant returns to scale Malmquist approach, this increases the number of linear programs from $N \times (3 \times t - 2)$ to $N \times (4 \times t - 2)$. However, Grifell-Tatjé and Lovell (1995) demonstrated that the Malmquist index does not in general correctly measure changes when variable returns to scale technology is assumed. Several authors have recommended the use of constant returns to scale specifications to avoid these problems (e.g. Coelli *et al.*, 1998: 228).

To apply the Malmquist DEA to the data set of this study, a minimum of 1159 linear programs had to be solved for a single constant returns to scale Malmquist DEA run. The software package DEAP was used. Table 7.4

summarizes the results obtained from these optimization runs with the 1991 to 1997 data.

It should be noted that the first year of observation is used to initialize the Malmquist indices and is therefore set to 1. The cumulative indices of technical efficiency change, technical change and productivity change are shown in Fig. 7.3.

The decline in total factor productivity of Austrian accommodation providers between 1991 and 1996 can be seen in Fig. 7.3 (dotted lines). Although

Table 7.4. Malmquist index summary of annual means.

Year	Efficiency change[b]	Technical change[b]	Productivity change[b]
1991[a]	1.000	1.000	1.000
1992	1.006	1.018	1.025
1993	0.967	1.004	0.971
1994	0.991	0.993	0.984
1995	1.013	0.993	1.006
1996	1.012	0.959	0.971
1997	0.989	1.034	1.023

[a] 1991 set to unity; [b]all Malmquist index averages are geometric means.

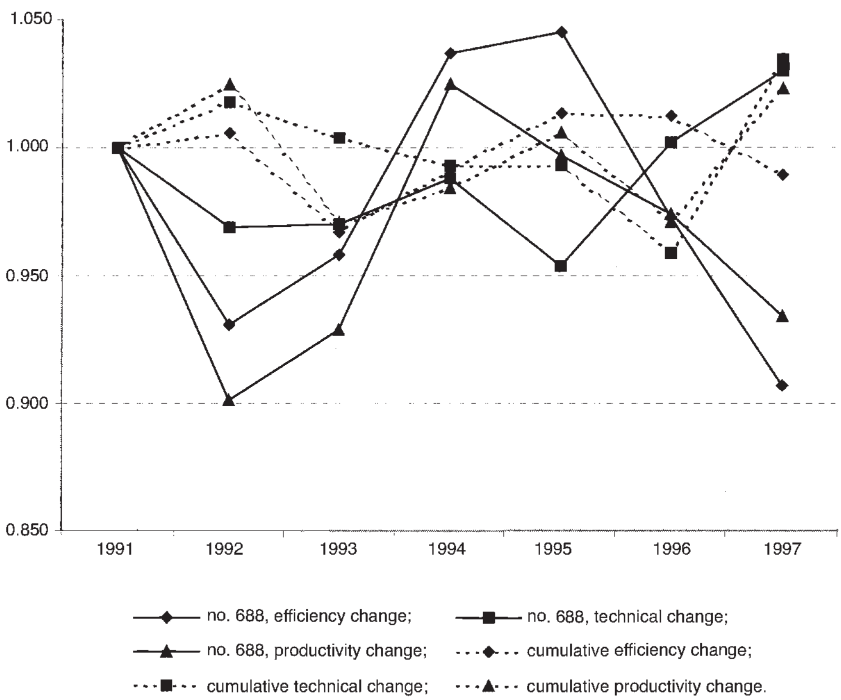


Fig. 7.3. Comparing the performance of hotel no. 688 with industry performance.

these results are based on a relatively small experimental data set, the negative developments in the Austrian hotel industry during 1991 and 1996 described earlier (see pp. 81–82) are observed. In the Malmquist analysis the total factor productivity index is decomposed to efficiency and technical change indices. The favourable course of the efficiency change index between 1993 and 1996 indicates the efforts of the industry to balance the drop in demand with improved efficiency (e.g. profit improvement programmes).

Malmquist indices for individual hotels can be used to assist hotel managers in benchmarking their performance against the decomposed industry performance indicators. A complete listing of the Malmquist indices for all companies in the database is given in Appendix Table A.8. For an example consider Fig. 7.3 in which the solid lines represent the performance of company no. 688. Overall, the manager of hotel no. 688 experienced unfavourable conditions between 1991 and 1997. In spite of 1994, the productivity change index is always clearly below the cumulative index, which indicates a poor productivity compared to the performance of the industry.

The decomposition of the productivity index allows multiple insights. One can assess whether the bad overall performance refers solely to inefficiencies in operations or involves poor anticipation of technological changes in the industry by the management of the company. Comparing the development of the technical change index with the cumulative function reveals that hotel no. 688 anticipated technological changes fairly well, thus leaving the primary explanation of the poor performance to the relative inefficiencies, which especially occurred between the periods 1991–1993 and 1996–1997.

The Malmquist extension to DEA has several merits compared to the window analysis approach demonstrated in the previous section. The major benefit is that it permits total factor productivity to be decomposed into technical change and technical efficiency change. It also has methodological merits; it neither requires the parameterization of a window width nor does it have idiosyncratic characteristics like the multiple representation of one and the same company in each window.

The disadvantage of the Malmquist extension is that there are no peer groups which derive automatically from the analysis as is possible with traditional DEAs. The dynamic version of the efficiency score, the efficiency index, is defined by two vectors of λs, one for the efficiency score evaluation in period t, and one for the efficiency score evaluation in $t + 1$. It is still unclear from the literature how these two vectors can be combined to form one common reference vector for all periods of observation.

7.1.3 Incorporating non-discretionary input factors

In business performance studies noise in the data set is very often caused by variations in the environment the companies are located in. There are two different approaches dealing with this noise in the data set: one is clustering the subjects into more homogeneously defined subgroups and performing the analysis on each of the clusters separately. The other option is the inclusion of background (or moderator) variables.

There are several performance studies in the literature where DEA has been applied iteratively to a predefined number of subject clusters. In a study of education in California, for instance, Sengupta and Sfeir (1986) split high-school districts into rural and urban; similar adjustments were made by Grosskopf and Valdmanis (1987) in an evaluation of hospital performance. In Ganley and Cubbin (1992) clusters based on recognized administrative groups of local education authorities largely amount to a distinction between rural and urban schooling.

An objection that typically arises in conjunction with clustering subjects in performance studies relates to the criteria chosen for defining clusters. Ganley and Cubbin express this problem when they say:

> Generally speaking, empirical clustering criteria have been rather crude and cannot exclude the possibility that a peer drawn from the same cluster may nevertheless be quite unlike the inefficient DMU [decision making unit[11]] for which it has been chosen. This will always be true in a trivial sense because every DMU is likely to have some unique characteristics (e.g. location).
> (Ganley and Cubbin, 1992: 135)

Ganley and Cubbin also argue that, notwithstanding this problem, the effects of clustering on efficiency remain valuable in clarifying the discriminating power of DEA in terms of its ability to identify meaningful targets and peer groups.

In response to this criticism, more rigorous procedures for clustering have begun to emerge in DEA literature. For example, Banker *et al.* (1989) developed an F-statistic that allows a test of the internal homogeneity. They tested this approach on a sample of 111 government-supported hospitals in North Carolina. They concluded that clustering can embody attainable targets by setting more demanding 'tight' targets for higher-performing groups and 'looser' targets (e.g. 90% of best-practice attainments) for the remaining group. They argue that this procedure, by making targets more equitable, will assist in acceptance of DEA assessment.

Another approach to deal with non-discretionary effects in DEA is the inclusion of moderator variables into the model. Although the inclusion of background variables is widely recognized in efficiency studies, there remains some dispute in the literature as to the effects of non-controllable variables on efficiency (e.g. Banker and Morey, 1986a,b; Golany, 1988; Ray, 1988,

---

[11] In 1981, Charnes *et al.* introduced the generic term 'Decision Making Units' (DMUs) to describe the collection of departments, divisions or administrative units that have common inputs and outputs and are being assessed for efficiency. In the following, this expression has been used by many authors on studies on productivity and efficiency analysis. In this book the author decided to use the terms firm, company and business interchangeably instead of the artificial term DMU.

1991; Ruggiero, 1996, 1998). A DEA model which incorporated the effects of uncontrollable inputs where reductions cannot be achieved was introduced by Charnes and Cooper (1985). The theoretical foundation and mathematical treatment by linear programming was discussed on pp. 71–73. In the following the effects of including environmental input variables are tested on the AHRP data set.

As stated earlier, initially, the inclusion of non-discretionary variables was omitted. Now such variables are defined by denoting X_1–X_3 as being uncontrollable to the manager. Recall that X_1 is the total number of beds, X_2 is the total number of seats in the F&B area of the hotel and X_3 is the number of opening days. The key to understanding which input variable is defined as discretionary and which is defined as non-discretionary lies in the observation that information about the extent to which a non-discretionary input variable may be reduced is not meaningful for the hotel manager. Note that this decision may not be the same for each manager for a given data set and that this may vary according to the manager's operational flexibility (which is again constrained by the environment).

Analysis and results

As the special software package DEAP did not support the processing of mixtures of controllable and uncontrollable variables, in the following, DEA problems were solved with Holger Scheel's Efficiency Measurement System (EMS), which uses Csaba Mészáros's state-of-the-art BPMPD linear program solver.[12] This allows one to efficiently estimate various forms of DEA models (input/output-oriented; radial/additive distance; constant/variable/non-increasing/non-decreasing returns to scale) and offers a variety of additional options (non-discretionary variables; calculation of Andersen and Petersen's 'superefficiency' scores; weight restrictions).

For the AHRP application an input-oriented variable returns to scale model was used. The complete results of the DEA runs with non-discretionary and discretionary inputs are listed in Appendix Table A.9. In comparison to Fig. 7.1, Fig. 7.4 shows the operating profit plotted against the DEA efficiency score distinguishing between discretionary and non-discretionary input variables for each of the 61 hotels in 1997. Again, the hotels deemed efficient and given a score of 1 are plotted along the right border. Comparing this plot with the plot in Fig. 7.1 clearly shows much higher inefficiencies for hotels which were run unprofitably or even experienced a loss in gross profit. In general, a strong positive correlation between gross profit and efficiency scores can be observed from the pictorial presentation in Fig. 7.1.

Hotel no. 2223, which is an 87-bed property, was selected for more detailed analysis and interpretation. In the year under consideration hotel no. 2223 achieved €978,500 F&B revenue and €328,600 room revenue, and

[12] www.sztaki.hu/~meszaros/bpmpd/

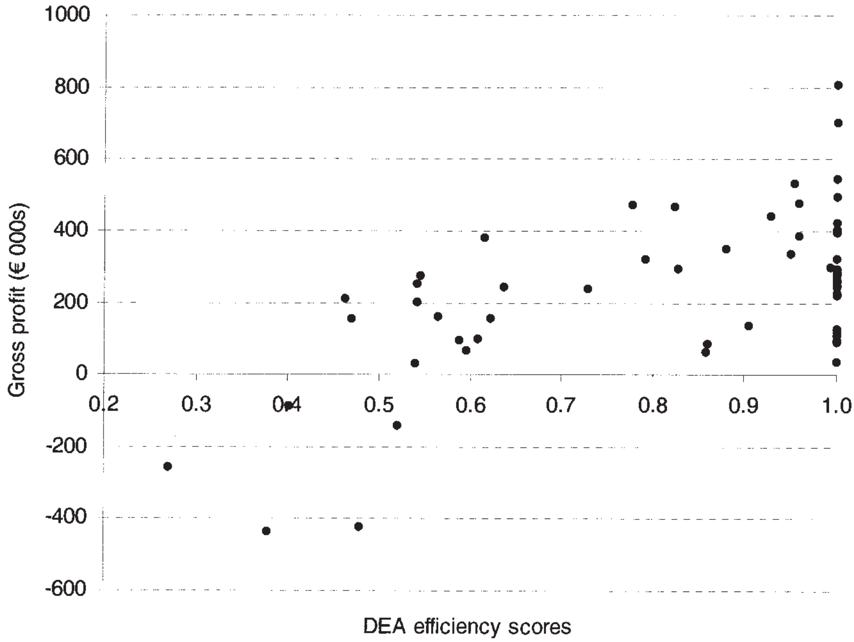


Fig. 7.4. Profitability versus DEA efficiency with non-discretionary input variables for 61 Austrian hotels in 1997.

spent €958,500 for operating expenses. The gross profit was €348,700, or 27% of the total revenue. The occupancy rate was 50.9%.

By using DEA with uncontrollable input factors it was found that four of the remaining 60 hotels (hotels no. 1061, no. 1353, no. 2336 and no. 2786) constitute a peer group for the company under evaluation. However, as is typically the case, the characteristics of the peer-group members do not perfectly match those of the unit being evaluated.

To account for the difficulty in comparing actual property-operation characteristics, the linear program builds a weighted composite of the four efficient peer-group members identified that perfectly match the levels of outputs and the operating environment of the unit being evaluated. Thus a composite, efficient benchmark hotel is derived to develop what may now be called a ‘scorecard’. It shows the resource-expenditure targets for hotel no. 2223 based on the management achieving at least the same total room revenues, F&B revenues, and occupancy rate in the same or an even more difficult environment. With the benchmarking figures listed in Table 7.5, it is possible to calculate the efficiency score for hotel no. 2223. Dividing the total expenditures of the composite benchmarking partners by hotel no. 2223’s expenditures results in the efficiency score of 0.880.

The peer group is matched to hotel no. 2223 based both on the non-controllable factors and on the levels of all outputs in so far as the

Table 7.5. DEA benchmarks for hotel no. 2223.

t = 1997	m = 35 (no. 2223)	m = 20 (no. 1061)	m = 27 (no. 1353)	m = 40 (no. 2336)	m = 48 (no. 2785)	BM[b]
Number of beds	87	130	81	58	65	84
Number of seats	120	27	210	85	140	120
Number of opening days	365	265	240	298	365	308
Total F&B revenue (€)	978,500	1,053,800	826,300	306,500	1,126,800	978,900
Total room revenue (€)	328,600	712,300	369,300	213,400	151,700	347,600
Occupancy[a] (%)	50.9	63.1	62.4	66.8	35.8	50.9
Total expenditures (€)	958,500	1,222,200	492,600	296,200	881,500	843,700
Payroll and related (€)	475,200	624,000	194,900	126,800	475,600	428,900
Material-type exp. (€)	212,700	249,100	105,500	82,900	236,000	200,400
Energy costs (€)	66,100	73,000	40,800	24,900	52,300	53,100
Cleaning costs (€)	13,600	22,100	12,900	4,700	22,700	19,000
Maintenance costs (€)	82,900	118,500	30,500	21,200	43,200	58,600
Communication costs (€)	16,900	18,000	9,200	7,800	7,800	10,700
Marketing costs (€)	58,700	46,600	45,800	7,800	2,800	23,400
Administration costs (€)	32,300	70,900	53,100	19,900	41,100	49,600
Gross profit (€)	348,700	543,900	703,000	223,700	396,900	482,900

[a]Per opening days; [b]benchmark (composite group) based on LP solution with $\lambda_{20}^* = 0.262$, $\lambda_{27}^* = 0.202$, $\lambda_{40}^* = 0.083$ and $\lambda_{48}^* = 0.453$; $e_{35}^* = 0.880$.

environmental factors in the benchmarking group are either equal or less favourable compared to the one of hotel no. 2223. In the case of the output factors, the virtual benchmarking company either performs equal to or better than the hotel under evaluation. The benchmark output figures are approximately the same as hotel no. 2223's figures, however, the operational expenditures are considerably less when compared with the hotel under evaluation.

The benchmarks that are most potentially interesting for management are those for the eight resource expenditures. They constitute a set of targets for hotel no. 2223 to work toward. If hotel no. 2223 utilizes its resources more efficiently, it should be able to achieve the same output with lower expenditures. The DEA analysis indicates that hotel no. 2223 has the potential to improve its output by €134,200 (improvement of about 38.5%), when it reduces it expenditures to €843,700 (a decrease of about 12%).

Figure 7.5 depicts the ideal direction for hotel no. 2223 to move to be more efficient and profitable, as well as showing the position of the four benchmarking partners as suggested earlier in a related context. The identity of the four peer properties can be made available to hotel no. 2223 so that the management can ascertain from them (perhaps through site visits) the details of the processes and practices that enable them to perform better.

7.1.4 Procedure for ranking efficient companies

One often-cited criticism of DEA is that it only allows dealing with inefficient companies. The concern is that very little, or even nothing (Ganley and Cubbin, 1992), can be said about the efficient companies in a data set. As far as active decision-making is concerned, the conclusion inferred is that DEA is only of use for suggesting adjustments to performance in relatively inefficient companies.

Tests by Ganley and Cubbin (1992) showed that clustered DEA tends to create an additional number of efficient companies in the results compared to a DEA with a pooled data set. Adding additional variables and hence increasing the number of constraints in the linear programs also increases the number of companies with identical unit-efficiency scores. More generally, research by Ahn and Seiford (1990) points to a link between the number of variables and the (average) efficiency of companies. In particular, as the size of the variable set increases the discriminating power of DEA declines with an increasing number of companies attaining best practice (Seiford and Thrall, 1990). This outcome is not so much a flaw in DEA, as a direct result of the dimensionality of the input–output space rising relative to the number of companies. An analogy may be drawn with the dimensionality constraints imposed by 'degrees of freedom' considerations in econometrics. The dimensionality problem may nevertheless diminish the managerial policy relevance of DEA. In particular,

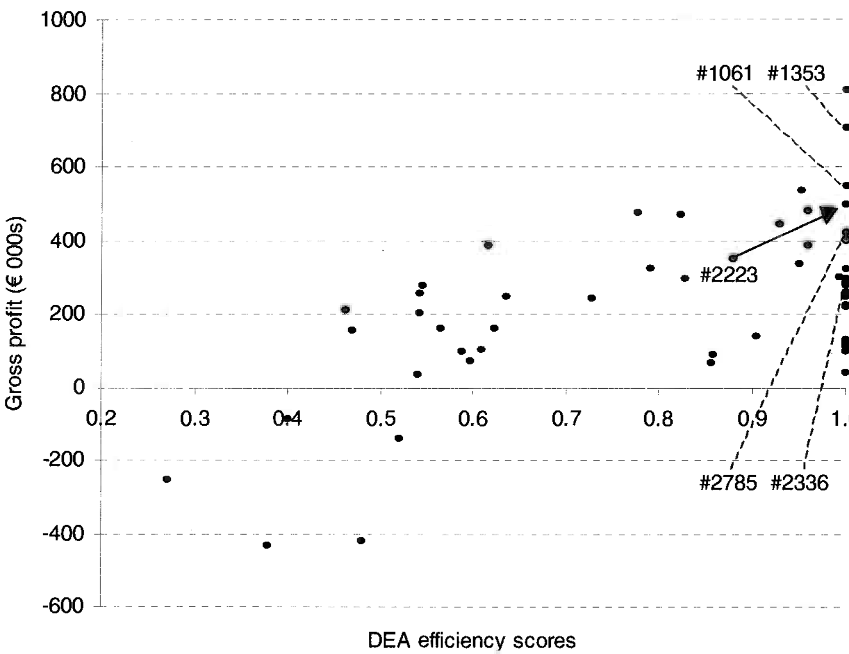


Fig. 7.5. Ideal position of hotel no. 2223.

real applications are constrained by the need to limit the number of companies labelled 'best practice'. Consequently, for the majority of companies in a data set nothing can be said about performance other than the implicit presumption that, in attaining best-practice, their performance is satisfactory and 'equivalent'. Since best-practice is not necessarily adequate in any absolute sense, it would be useful for the decision-maker to have more guidance on the quality of best-practice performance.

Andersen and Petersen (1993) suggested a modified version of DEA, which allows the ranking of efficient companies. Their basic idea was to compare the company under evaluation with a linear combination of all other companies in the sample, however, the company itself is excluded. Under this approach it is possible for an efficient company to increase its input vector proportionally while preserving efficiency. The company obtains, in that case, an efficiency score above one. The score reflects the radial distance from the company under evaluation to the production frontier estimated with that company excluded from the sample. The approach provides an efficiency rating of efficient companies similar to the rating of inefficient companies.

An illustration of Andersen and Petersen's (1993) idea is given by Fig. 7.6. Consider the evaluation of the efficient company P_3 in Fig. 7.6. According to the definition of the DEA efficiency measure, the reference point in the evaluation of P_3 is the observation itself, and P_3 is assigned the index one. Elimination of P_3 in the spanning of the reference set implies that P_3 is compared to that (inefficient) point in the input possibility set spanned by the remaining set of observations with the minimal distance of P_3. The reference point thus becomes P_3'. In analogy to the inefficiency index, the efficiency index is calculated by $0P_3'/0P_3$ and has the same interpretation as the Farrell measure: P_3 may increase its input vector proportionally up to the efficiency index and still remain efficient, but it will be dominated by a combination of P_2 and P_4 if the proportional increase in the input vector exceeds the efficiency score.

Although this technique by Andersen and Petersen is obviously an important extension and can enhance existing DEA models, it is rarely applied in

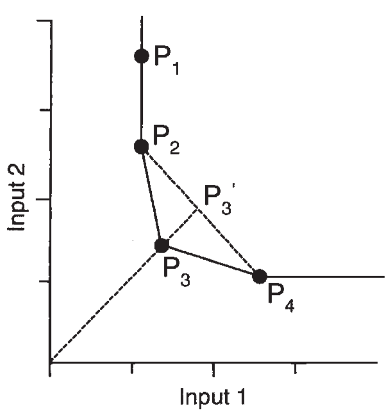


Fig. 7.6. Ranking efficient companies (Andersen and Petersen, 1993).

performance research applications. One of the few exceptions is Jammernegg *et al.* (1997).

In the present study, the Andersen and Petersen procedure for ranking efficient companies is applied to the previously achieved DEA results of discretionary and non-discretionary input factors. In this model, out of 427 evaluations, 208 (48.7%) hotels were identified as being efficient. Hence for almost 50% of the hotels in the database nothing else could be said as they performed equally efficiently.

Repeating the analysis with the Andersen and Petersen modifications for efficient companies resulted in a high number of infeasible solutions (59.1%). This is in line with results by others who have found that estimates for a considerable number of companies are undefined because of the infeasibility of the set of constraints of the modified DEA model (Pastor *et al.*, 1999). Boljunčić (1999) gives an explanation for this high number of infeasible solutions. He identifies that the main problems are caused either by zero values in the variable set or by cases where some companies show extremely high efficiency values (Boljunčić, 1999: 243). In general, infeasible solutions occur in attempts to solve large, complicated linear programming problems, where the constraints specified cannot be simultaneously satisfied (Schrage, 1997: 7). There is nothing much to do with infeasible linear programming solutions, hence the cases impacted must be excluded from further analysis.

The complete results are summarized in Appendix Table A.10. For an individual efficient company the analysis can provide two meaningful insights. First, a company can monitor its efficiency development even for years when the firm has been classified as a 'best-practice' company in the sense of Farrell. Figure 7.7 gives an example for hotel no. 14 which was a best-practice company in the years 1993, 1994, 1996 and 1997 (white bars) and an

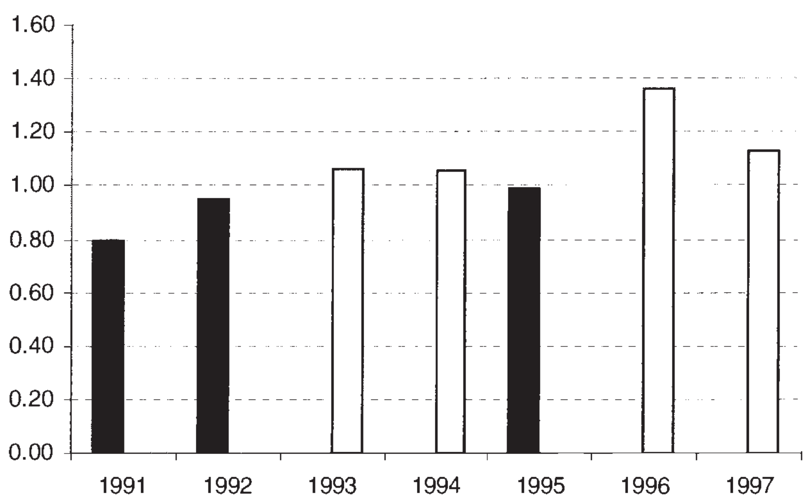


Fig. 7.7. Inefficiency (■) and efficiency (■) scores for hotel no. 14.

inefficiently operated company in the years 1991, 1992 and 1995 (black bars). The figure clearly indicates the very regular development, gradually increasing between 1991 and 1993, and declining slowly from 1993 to 1995. Note that hotel no. 14 crossed the 'efficiency line' twice, first from an inefficient to an efficient company and then back to an inefficient group of companies. 1996 was an extraordinary year for hotel no. 14, which certainly has to be considered in an overall evaluation.

There are several new procedures suggested in the literature for the analysis of efficient companies using DEA results. However, all of these extensions have in common that they introduce a complex form of two-stage approach in model formulation and computation (e.g. the 'slack adjusted DEA model' suggested by Sueyoshi *et al.* (1999); the DR/DEA model by Sinuany-Stern and Friedman (1998); or the 'single price system extension' by Ballestero (1999), which classifies efficient but not inefficient companies).

The model selected here is an input-oriented model, which seeks to identify technical inefficiency as a proportional reduction in input usage. As discussed in Chapter 4, it is also possible to measure technical efficiency as a proportional increase in output production, however, the former is the more adequate model for the manager. This is because hotel managers usually have objectives to fulfil, either set by corporate management goals or by self-defined business plans, and hence the input quantities appear to be the primary decision variables. In other applications it will also be possible that managers may be given a fixed quantity of resources and asked to produce as much output as possible. In this case an output orientation would be more appropriate.

7.2 Strengths and Weaknesses of the DEA Approach

DEA introduced new principles for making inferences from empirical data. Several studies have been performed in order to compare the capabilities of DEA with traditional multivariate methodologies, including regression analysis (Thanassoulis, 1993; Cubbin and Tzanidakis, 1998), discriminant analysis (Retzlaff-Roberts, 1996; Sueyoshi, 1999), artificial neural networks (Athanssopoulos and Curram, 1996) and even multiple criteria techniques such as the Analytical Hierarchy Process (Tone, 1989; see also the development of multiple-criteria DEAs by Shang and Sueyoshi, 1995, and Li and Reeves, 1999).

Recently the awareness of DEA for specifying performance benchmarks has increased: (i) for management science (Hawdon and Hodson, 1996; Morey and Morey, 1999; Post and Spronk, 1999; Sueyoshi *et al.*, 1999; Zenios *et al.*, 1999); (ii) for activity-based management (Kantor and Maital, 1999; Mota *et al.*, 1999); (iii) for portfolio management (Santos and Dyson, 1997: 6; Sarrico and Dyson, 1998: 4; Soteriou and Zenios, 1999); and (iv) for quality management (Mathiyalakan and Chung, 1996; Madu and Kuei, 1998). DEA enables efficient firms to be identified. It calculates target reductions in specific inputs for less-efficient firms and it enables tests with distribution-free or non-parametric procedures to be used to investigate the important factors contributing towards excellent performance for companies (Hawdon and Hodson, 1996). This acceptance of DEA is mainly because it focuses on observed operating practice, and circumvents specifying the complete functional form of the production function. It has greater practical appeal and higher perceived fairness than normative industrial engineering standards. Moreover, recent extensions to DEA have offered substantive flexibility to incorporate realistic assumptions, such as variable returns-to-scale properties and non-discretionary input variables.

Although DEA has several advantages over other methodologies for performance evaluation, it nonetheless suffers from a variety of weaknesses which should be subject for future research.

7.2.1 Stability of DEA results

DEA is not endowed with any formal system of hypothesis testing (Seiford and Thrall, 1990). This is because DEA is a non-statistical technique which makes no explicit assumptions on the distributions of the residuals. This, and other problems, have left DEA open to criticism. In recognition of these difficulties Charnes *et al*. (1985b) initiated work on a DEA-sensitivity analysis. Subsequently, research has begun to focus increasingly on sensitivity issues (Banker and Morey, 1989; Epstein and Henderson, 1989; Sengupta, 1990, 1992b,c; Banker, 1993; Banker *et al*., 1993, 1998; Retzlaff-Roberts and Morey, 1993; Hougaard, 1999; Maital and Vaninsky, 1999).

The initial work by Charnes *et al*. (1985b) only involved an examination of the effects on the efficiency score of deleting variables. However, if noise is present in an observation located inside the efficiency frontier, the consequences are limited to that company under evaluation. The inefficiency score of that company will be biased towards or away from the frontier depending on the nature of the distortion. The inefficiency scores of the other companies in the data set will not be affected. Data errors of companies which make up the frontier are more serious since these will change the efficiency score of all companies in the data set for which efficiency is defined by reference to the biased company.

An old suggestion to address this problem is given by Timmer (1971) who argued that a Farrell boundary can be constructed iteratively. In an iterative approach he successively eliminated outlying data points and re-estimated the frontier until the resulting efficiency estimates stabilized. This is possible in relatively large data sets, although it means that excluded units will have no efficiency score in the final iteration. The Timmer adjustment is arbitrary to the extent that it is not clear, *a priori*, precisely when the efficiency scores have stabilized to a sufficient degree to accept that random outcomes have been eliminated. Later developments of the Timmer approach can be found in Sengupta (1987, 1988), and Sengupta and Sfeir (1988).

More recently, Banker *et al.* (1993) performed extensive Monte Carlo simulations which suggest that the reliability of DEA results deteriorates considerably in comparison to an econometric approach when measurement errors become large. Retzlaff-Roberts and Morey (1993) introduce the concept of minimum frontier to allocative DEA in order to identify significantly inefficient units.

In a study reported by Ganley and Cubbin (1992: 128), a data error on one variable at one company reduced the average efficiency of the whole cross-section by 12%. Seven companies, formerly efficient on the correct data, achieved non-unit efficiency scores in the error-ridden data set. They argue that noise in outcomes might be identified in unexpected or abrupt change in the efficiency ranking of utilities year-by-year. In their study they suggest using Spearman's rank correlation coefficients to test whether the efficiency rankings change significantly when excluding individual companies from the data set. A high correlation represents stable efficiency scores which could then be the basis of acceptable targets.

Another approach was taken by Färe *et al.* (1987) who found more stable estimates when performing separate DEAs on successive cross-sections and deriving mean efficiency scores for companies. Also Brockett and Golany (1996) suggested that the analysis should be performed by group, rather than by individual units, which leads to stochastic extensions of DEA where random deviations from the group's behaviour can be studied.[13] A similar form of data pooling was suggested by Charnes *et al.* (1985a) as part of their 'window-analysis', which was discussed in detail on pp. 129–133. The resulting composite frontier derived by the window-analysis gives less weight to unusual observations and is therefore more robust to stochastic events.

Banker *et al.* (1998) present a stochastic data envelopment analysis (SDEA) model to estimate standards from comparative benchmarking data. The authors argue that their model can create mix and yield variance ratios when including estimates on substitutability or separability between factors. They illustrate their approach with data on nursing services from 66 state hospitals in one US state. Banker *et al.* (1998) show how one hospital's performance can be matched against the benchmark cost for the hospitals as a group, and conclude that SDEA sets more achievable standards than conventional DEA. Recently, however, it has been reported that the stochastic DEA models can only outperform the traditional DEA in some specific situations, but on average they cannot compete with the older techniques (Resti, 2000: 559).

Some valuable contributions have been made by Sengupta (1990) and Banker *et al.* (1989) who have examined methods for identifying 'gross data

---

[13] Their statistical evaluation of observed group differences was recently improved by Sueyoshi (1999).

errors' and regions of data stability. However, these developments are rather *ad hoc*. More promising extensions involve the incorporation of fuzzy mathematical programming which has been demonstrated by Sengupta (1992b,c) and Hougaard (1999). The ultimate research objective in order to distinguish accurately between measurement errors and inefficiencies probably lies in some form of marriage of parametric and non-parametric techniques. This would circumvent one of the principal difficulties currently inherent in a non-parametric frontier approach.

Another goal of DEA simulation is to impose control on allowable solutions, as in real world applications not all factors can be controlled by the managers. The cone-ratio DEA model (Charnes *et al.*, 1989, 1990) and the assurance-region aspects (Thompson *et al.*, 1986, 1990) are examples where upper and lower bounds are imposed on the weights to assure that certain environmental considerations and expert opinion are incorporated into evaluation. Extensions to these models have been introduced by Kao (1994) and Cooper *et al.* (1999).

7.2.2 Interactive DEA

Several advantages result when DEA is incorporated in an interactive environment. For example, the input and output variables must be carefully selected to make the analysis useful for the manager. Although DEA has fewer limitations than other econometric approaches in the choice of input and output variables, formal selection criteria are unavailable, the input–output variables in a model are therefore usually selected based on intuitive or pragmatic considerations (Haag *et al.*, 1992). Roll *et al.* (1989) attempt to give some guidelines on selecting the appropriate variable set for a DEA.

The advantage of an interactive system is that the user can go back and forth and learn from the output. The manager can change the variables selected and he/she is not bound to a strict classification as is usual in ordinary printed publications of panel studies. Hence the user will soon realize that results may vary significantly, sometimes even through minor changes in the variables selected. The manager who manipulates options can gain more insights and a better understanding of how to interpret benchmarking results and how to use them for managerial purposes.

Furthermore, this simulation environment may be the ideal platform for the implementation of more dynamic DEA models, like the 'inverse DEA model' recently introduced by Wei *et al.* (2000). Their extension of the basic DEA model solves such problems where, for instance, among a group of companies, a particular company's inputs are increased and, by assuming that this company maintains its current efficiency level, the maximum achievable increase in outputs is calculated. The inverse DEA model proposed by Wei *et al.* is certainly a useful tool to perform what-if analyses, which are perfectly suited for being integrated in an interactive decision support system.

An appropriate interactive decision support system is best implemented in a multi-user environment like the Internet in that the widest possible group could benefit. Extranet applications, like TourMIS, could offer online databases of financial and non-financial hotel data for DEA analysis, especially for SMEs which are less organized in the exchange of business data than international hotel chains. Finally, the advantage of a real-time application is that additional insights can be gained by multi-period analysis and extrapolations of business data time series. An appropriate linkage of a DEA to a database system can therefore easily convert usage of DEA model from an *ex post* evaluation instrument to a prospective oriented instrument which might also support budgeting tasks for small and medium-sized enterprises.

However, there are severe problems which have to be addressed when developing a real-time interactive DEA system. For example, the standard DEA model is a static, one-period evaluation and difficult to integrate in an interactive environment. When a new company's data are added to a database they become a part of one or more subsets of the data in which its presence must be considered. For these subsets current methodology means that all efficiency runs for all other firms have to be repeated to redefine the efficiency frontier. It would clearly be appropriate to move towards more general dynamic DEA modelling in order to handle trend data in growing organizations and changing environments. A step in support of this would be deriving an explicit partial adjustment mechanism to use rather than replicating all linear programming runs.

Chapter 8

Evolutionary Data Envelopment Analysis

8.1 Motivation for an Evolutionary Data Envelopment Approach

Data Envelopment Analysis (DEA) has been discussed and applied extensively in research and practice during the last two decades. Recently, user-friendly computer programs have been developed which address conceptual and methodological problems in using DEA. These software packages offer the convenience of automatically solving all linear programs which are necessary to calculate efficiency scores for all companies in a data set. Additionally, DEA software packages sometimes offer statistics like detailed slack analysis or diagnostics of decreasing or increasing returns to scale.

The linear nature of the DEA problems, so far being proposed in research and practice, makes the simplex algorithm a feasible routine. The calculations in DEA software packages are performed by means of the simplex search methodology. The simplex search method uses a simplex (the feasible region) of solutions to create a new and better simplex according to some rules, depending on the objective function values at all solutions of the simplex. The essential idea is to generate a search direction using the simplex and to follow this direction until the optimal solution is found. Since the evaluation function is linear, the optimal solution lies at one of the vertices of the simplex or, in a degenerate case, anywhere along a boundary.

The simplex algorithm for linear optimization problems, or problems that can be transformed into linear form, is, in principle, well defined and leads, in most cases, to unambiguous optimal solutions. However, the simplex approach has several limitations which may obstruct the development of DEA in the future. First, the underlying linear programming approach only applies directly to situations in which the constraints are linearly defined. The linearity requirement features proportionality, additivity and continuity of the variables in the DEA model.

However, a model that includes the two input variables 'resource costs per unit sold' and 'quantity of resources used' is probably not linear. In fact, business processes by their nature are highly complex, open systems, which interact both with other processes within the same firm and within their environment (suppliers, customers, labour market, etc.). Competition for scarce resources between different processes results in a non-linear interdependence between interacting components over time (Shiroma, 1996).

The proportionality requirement, or the effect of a single variable or activity by itself being proportional, is satisfied, but the interaction between the two input variables is multiplicative rather than additive. As it has been shown, there exist extensions to include returns to scale assumptions in DEA models in order to overcome some of the proportionality violations which may occur in DEA models. However, a model that includes the decision variable 'number of hotel rooms to build' might satisfy the proportionality and additivity requirement but may violate the continuity conditions. The recommendation having or not having a spa, restaurant or private beach might be difficult to implement. Once the evaluation function is non-linear, multimodal or even discontinuous, the basic linear programming methodology breaks down (Michalewicz and Fogel, 2000).

Another problem with linear programming is when the formulator has been too demanding and an infeasible solution occurs. As seen from the above discussions on the calculation of DEA scores for efficient companies, not all DEA solutions are necessarily robust. The non-existence of a feasible solution depends solely on the constraints of the DEA. Infeasible solutions occur in large complicated problems and when trying to solve unbounded problems. In the latter case concluding the formulation by saying that an infinite amount of profit can be made is unrealistic. However, in large problems there are typically several variables that are unbounded, and it is not easy to identify the manner in which this phenomenon arises.

A third problem within linear programming is that sometimes there are several optimal solutions with unique objective function values. Such degenerated solutions occur in small examples when redundant constraints exist (Schrage, 1997: 16). However, they may also occur in situations when a company rests on the efficiency frontier, but is still subject to further improvement (see the 'slack optimization problem' displayed in Fig. 4.8).

Linear programming has a very long history and there has been much work done and many extensions implemented that have been useful in fixing most of these problems in practice. Research on computational aspects of DEA has reflected these advances in order to create highly efficient, specialized codes (Phillips *et al.*, 1990; Ali, 1992, 1994; Ali and Seiford, 1993; Charnes *et al.*, 1993; Barr and Durchholz, 1997). On the other hand, these extensions make DEA applications very problem specific and less flexible. For instance, in a real-time environment the simplex algorithm occurs as relatively inflexible and deterministic which can hinder automatic adaptations (e.g. in model design). The simplex approach also prevents the detection of (non-optimal) alternative sets of comparison partners, which, nevertheless, can be significantly better than the company under evaluation.

8.1.1 Multi-modal characteristics of DEA problems

The DEA analyses in Chapter 7 have raised the impression that there is only one optimal solution which delivers one set of comparison partners which can be used for efficiency analysis, however, this is not completely true. In fact, DEA programs usually find only one solution out of many solutions which cannot be improved in a sense that any change in a weight will lead to a better solution.

To illustrate this, consider the following simple example of a constant returns of scale input-oriented DEA with five observations, one output and two input variables (the example is taken from Coelli, 1996: 17) for company P_4 shown in Fig. 8.1.

The optimal solution derived by standard DEA software results in company P_4 being inefficient with a DEA score of $e_4 = 0.714$. Concerning the optimal selection of comparison partners the DEA algorithm suggests companies P_2 and P_5 as peer members with the weightings $\lambda_2 = 0.214$ and $\lambda_5 = 0.286$, respectively. All other companies are not considered for any comparisons.

That this result is an optimal solution is proven when evaluating the mobility of each company's weightings (see Table 8.1). The lower bounds in the present example have reached the final solution of weightings for company P_4, which indicates that they cannot be lowered any further. Due to the fact that this simple model does not include non-discretionary variables, the upper boundaries are all set as infinite. However, any changes in each of the

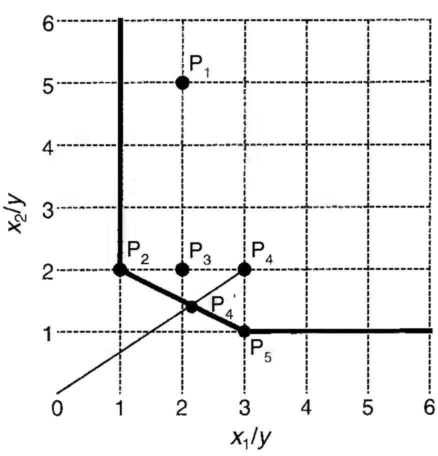


Fig. 8.1. Simple input-oriented DEA example (Coelli, 1996: 17).

weightings towards the upper bounds would be unfavourable as this would deteriorate the DEA score. Thus, the result displayed in Table 8.2 is definitely an optimal solution.

Now consider the situation illustrated in Table 8.3. In this example companies P_3 and P_5 have been identified as peer members of company P_4, which is an obvious solution when company P_2 is omitted from the analysis (see Table 8.2). The DEA score e_4 calculated with this set of weightings is 0.8, which is also below 1 and therefore indicating inefficiencies of company P_4.

This solution is also an optimal solution as it cannot be improved by lowering any weightings without violating the output constraints, nor can it be improved by raising any of the weightings without worsening the DEA inefficiency score. In order to get to the global optima illustrated in Table 8.2, at least two company weightings must be changed simultaneously.

Now consider the situation illustrated in Table 8.4. In this example also, company P_3 has been identified as a peer member although, as can be seen from Fig. 8.1, company P_3 is not even situated on the frontier. The DEA score

Table 8.1. Simple input-oriented DEA example (Coelli, 1996: 17).

Company	y	x_1	x_2	x_1/y	x_2/y
1	1	2	5	2	5
2	2	2	4	1	2
3	3	6	6	2	2
4	1	3	2	3	2
5	2	6	2	3	1

Table 8.2. Global optimum in the simple DEA example.

Company	λ	λ_l	λ_u
1	0	0	∞
2	0.214	0.214	∞
3	0	0	∞
4 (e = 0.714)	0	0	∞
5	0.286	0.286	∞

Table 8.3. A local optimum in the simple DEA example.

Company	λ	λ_l	λ_u
1	0	0	∞
2	0	0	∞
3	0.2	0.2	∞
4 (e = 0.800)	0	0	∞
5	0.2	0.2	∞

Table 8.4. Another local optimum in the simple DEA example.

Company	λ	λ_l	λ_u
1	0	0	∞
2	0.100	0.100	∞
3	0.167	0.167	∞
4 (e = 0.850)	0	0	∞
5	0.150	0.150	∞

calculated with this set of weightings is 0.850, which is lower than in the previous example, but still classifies company P_4 as being inefficient.

This solution is also an optimal solution in the previous sense, but a remarkable difference here is that additionally it interrupts the convex shape of the efficiency frontier.

The important issue stressed here is that for the optimal selection of comparison partners it might not only be important to locate the global optimum which is found with traditional DEA applications. The conventional DEA does not give any answer to the question: what are other good sets of 'best practising' partners? Maybe there are alternatives which get on with a completely different set of comparison partners and only a small loss in efficiency gain compared to the global optimum. It might be rather important to locate all optima above a certain threshold, e.g. situations where $e < 1$, than to inspect only one absolute optimal solution.

For example, in Table 8.4 company P_3 was identified as a peer partner with the highest weighting compared to the other peers in the set of 'best practising' partners. This is certainly important information for the manager of the company under evaluation, especially when there are obstacles which arise in following a benchmarking process because one or more preferred comparison partners turn out to be unsuited. Accepting the multiple possibility concept it is certainly also important for softening the deterministic and immovable perspective on solutions fostered by DEA algorithms.

For completeness it must be mentioned here that the above example can easily be extended for DEA problems with non-discretionary input variables. In this case, the global and locally optimal sets of company weightings have either upper or lower bound values and any changes within the movable range will deteriorate the DEA inefficiency score.

In summary, the traditional linear programming optimization approach for DEA is appropriate for well-behaved, unimodal, simple objective functions. When applied to multimodal problems where good solutions, not necessarily the best solution, are of interest, this method either is not very efficient or cannot be used at all. For this reason the author proposes an alternative method to be used for solving DEA problems, especially when the objective is to find alternative (sets of) benchmarking partners.

8.1.2 Optimization techniques in tourism research

There is no intention to imply that traditional linear programming algorithms are useless, in fact they have been used extensively in many engineering optimization problems. The following sections describe a robust search and optimization method which works very differently from the traditional methods and which has been successfully applied to solve a wide variety of search and optimization problems in sciences, engineering, and commerce. The suggestion here is that a different approach sometimes leads to new ideas and possibilities that may become beneficial for the underlying problem, the efficiency measurement and the selection of appropriate comparison partners.

Traditional methods

Although optimization methods are widely used in management science (Anderson *et al.*, 1997), they are seldom adopted by the tourism research community. In general, there are only a few applications of optimization techniques in the tourism field (Mazanec, 1986a,b; Kottke, 1988; Van der Knijff and Oosterhaven, 1990; Canestrelli and Costa, 1991; Hadjinicola and Panayi, 1997; Taplin and McGinley, 2000).

Traditional optimization techniques include Canestrelli and Costa's (1991) study applying linear programming models to determine the optimum level of tourism carrying capacity in urban destinations. Their model is made operational within a fuzzy linear programming approach that is tested in a case of the historical centre of Venice. Kottke (1988) and, later, Van der Knijff and Oosterhaven (1990) also use linear programming to determine the optimum level of tourism development in an area and to derive the optimum combination of government policy tools necessary to maximize tourism employment in a region. Mazanec (1986a,b) introduced a dynamic programming approach in the final step of a more complex budget allocation model tailor-made for national and regional tourism offices. Hadjinicola and Panayi (1997) address the general problem of a hotel's need to overbook accommodation in order to maximize capacity. They developed a mathematical model for the determination of minimum expected loss and demonstrate that the minimum expected loss is smaller when treated at the hotel level rather than at the level of the individual tour operator. Recently, Taplin and McGinley (2000) used a multiperiod linear programming model to simulate daily car touring decisions, which they have observed from Australian tourists.

However, the major field of application of optimization techniques in tourism research is the airline industry. Areas where optimization techniques have been introduced include personnel scheduling, ground holiday delays, airline network design, flight frequency determination and optimization problems in yield management systems (see Hurley and Moutinho (1996) and Hurley *et al.* (1998) for a detailed list of references). Linear programming is the most popular technique in these studies.

Evolutionary algorithms

Applications of evolutionary algorithms in management science are covered by numerous researchers. However, there has been a considerable bias towards scheduling problems. Evolutionary algorithms proved to be quite successful in finding good solutions to complex problems involving the travelling salesman, mass scheduling and graph partitioning, but also for engineering problems like the design of bridge structures or the optimal use of power plants. Nissen (1995) provides a comprehensive list of more than 500 references of evolutionary algorithm applications in management science. Considering the applications of evolutionary optimization methods in tourism research the number of applications is rapidly declining. Only a few attempts have been made to adopt evolutionary optimization methodologies and these concentrate almost exclusively on the application of genetic algorithms (Schifferl, 1996, 1998; Taplin and Qiu, 1997; Hurley *et al*., 1998).

Taplin and Qiu (1997) used a genetic algorithm (GA) to simultaneously estimate a gravity model of trip generation and a route assignment model for car tourists on their way to a destination and back home. Their GA was a simplified version of the GENOCOP (Genetic Algorithm for Numerical Optimization of Constrained Problems) system proposed by Michalewicz (1996), which is notable as this is the same basis for the system presented here. Taplin and Qiu's application was remarkably the only GA using floating-point presentations in tourism research.

Hurley *et al*. (1998) discuss the application of genetic algorithms for the problem of tourism site location. Their results suggest that genetic algorithms are likely to outperform traditional optimization methods when the number of sites (existing or proposed) is large.

Schifferl (1996, 1998) introduces an application of genetic algorithms for product design optimization. In her work she applies genetic algorithms to discover optimal multi-attributed products for different customer segments on a case study of holiday homes in former log cabins in alpine areas.

8.1.3 Other AI approaches for the selection of optimal comparison partners

There have been a few other attempts at applying artificial intelligence methodologies for the selection of optimal comparison partners in the literature. They are mentioned here for completeness, but will not be discussed in detail.

The objective of a study by Back *et al*. (1995) was to investigate the potential of neural networks for preprocessing large amounts of financial data, and for presenting the approximated financial performance position of one company as compared to that of others. In order to do competitive benchmarking they used self-organizing maps on a database of international pulp and paper companies. Although they describe their findings as encouraging, their clustering of companies with similar financial characteristics did not solve the stated problem. In fact, the authors did not explicitly address how efficiency measurement should be performed in their results.

Neural networks were used to assess the performance of companies against an expected performance in the example data set which has been made available with a well-known neural network software package.[1] Thus, a comparison of DEA and artificial neural networks was provided on a set of bank branches by Athanassopoulos and Curram (1996) as a tool for assessing the efficiency of companies. The authors stress the advantages of neural networks compared to ordinary regression analysis, but due to inherent differences between DEA and neural networks, they propose the complementary use of these alternative methods for assessing performance.

8.2 The Principal Idea of a Genetic Algorithm

Here the principal idea of a genetic algorithm (GA) and alternative approaches are discussed only to an extent that is required by what is presented subsequently. Although evolutionary computation is a relatively new area, there exists a vast literature on this subject. Readers interested in the origins of evolutionary computation are advised to refer to Darwin (1859, 1985), Box (1957), Holland (1962, 1975), Rechenberg (1993), De Jong (1975) and Schwefel (1994).

GAs were initially developed by Holland (1975). They provide a search and optimization procedure that is motivated by the principles of natural genetics and natural selection. Some fundamental ideas of genetics are borrowed from the genetic processes of biological organisms. Darwin's principle of natural selection ('reproduction and survival of the fittest') and its adaptation by GAs made this procedure very famous. Although this slogan seems to be slightly tautological in the natural environment (Dawid, 1999), it is very useful in optimization problems, where the fitness is defined as the value of a function to be optimized.

The working principles of GAs are very different from those of most traditional optimization techniques. They transform a population of individual objects, each with an associated fitness value, into a new generation of the population using the Darwinian principle and analogies of naturally occurring genetic operations, such as crossover (sexual recombination) and mutation. Other biological phenomena, such as the duality of genes or the existence of dominant and recessive genes, are up to now (2001) not considered by evolutionary algorithms. However, the importance of these processes for evolution in nature is still not completely known in a biological sense and

[1] 4Thought, Right Information Systems Ltd, London.

the choice of the selected features that are transferred to GAs is due mainly to implementation problems (Dawid, 1999).

A genetic algorithm is an iterative procedure that operates on a constant-sized population of individuals, each one represented by a finite string of symbols, known as the chromosomes, encoding a possible solution in a given problem space. This space, referred to as the search space, comprises all possible solutions to the problem at hand. Each individual in the population represents a possible solution in this search space. The genetic algorithm attempts to find a satisfactory solution to the problem by genetically breeding the population of individuals over a series of many generations. Generally speaking, the genetic algorithm is applied to spaces which are too large to be exhaustively searched.

An introductory overview on evolutionary computation including genetic algorithms, evolutionary programming, evolution strategies, genetic classifier systems and genetic programming is provided by Bäck and Schwefel (1993) and Koza (1997). Concerning GA, John Holland's pioneering book *Adaptation in Natural and Artificial Systems* (1975) showed how the evolutionary process can be applied to solve a wide variety of problems using a highly parallel technique. Until today, though, the main textbook reference for many GA researchers is David Goldberg's *Genetic Algorithms in Search, Optimisation, and Machine Learning* (1989). Additional textbook information on GAs can be found in Koza (1992), Kinnebrock (1994), Mitchell (1996), Michalewicz (1996), Banzhaf *et al.* (1998), Dawid (1999) and Michalewicz and Fogel (2000).

8.2.1 GA approaches for constraint handling

Several approaches have been proposed for solving general non-linear programming problems through GAs (Powell and Skolnick, 1993; Joines and Houck, 1994; Michalewicz and Schönauer, 1996). Most of them are based on the concept of penalty functions, which penalize infeasible solutions, for handling non-linear programming problems. Although several ideas have been proposed about the design of the penalty function, this method has several drawbacks which led to disappointing results in several experiments as pointed out by Michalewicz and Schönauer (1996), Michalewicz (1996) and Michalewicz and Fogel (2000).

The EDEA design is a revised version of Michalewicz's GENOCOP system, a GA for solving general linear programming problems by avoiding the drawbacks of the penalty methodology. One of the distinct characteristics of GENOCOP is its floating value representation of the chromosomes. One of the drawbacks of traditional GAs using a (binary) coding scheme is that a proper coding of the problem to be addressed needs to be used. When using GENOCOP to solve multiparameter optimization problems, this is not needed as in this algorithm a string is composed of a set of real values. DEA parameters are numeric, so representing them directly as numbers, rather than bit-strings, seems obvious and may have advantages. Janikow and Michalewitz (1991) made a direct comparison between binary and floating-point representations, and found that the floating-point version gave faster, more consistent and more accurate results.[2]

With GENOCOP, Michalewicz was the first to show that the floating point representation in a GA implementation can be faster, more consistent from run to run and provides a higher precision specifically in large domains where binary coding is rather inefficient.

In 1992, when Michalewicz introduced his original system, he also introduced a general non-linear programming version, GENOCOP II. Later, in GENOCOP III he incorporated the original GENOCOP system for linear constraints, but extended it by maintaining two separate populations, where a development in one population influences evaluations of individuals in the other population (Michalewicz and Nazhiyath, 1995; Michalewicz and Schönauer, 1996). GENOCOP III was recently tested by Sakawa and Yauchi (1999) for multiobjective non-convex programming problems. Taplin and Qiu (1997) were the first to apply the original GENOCOP system to a tourism research application.

8.2.2 GA extensions for multiobjective problems

The Vector Evaluated Genetic Algorithm (VEGA), an early GA application on multiobjective optimization by Schaffer (1984, 1985), opened new possibilities of research in this field. In his work, Schaffer tried to capture all Pareto optimal solutions of a multiobjective optimization problem. His main idea was to divide the population into equal-sized subpopulations, each subpopulation responsible for a single objective. The selection procedure was performed independently for each objective, but crossover was performed across subpopulation boundaries. Additional heuristics were developed to decrease a tendency of the system to converge towards individuals which were not best with respect to any of the objectives.

This new special research area within GA was extensively discussed and extended by many other researchers (e.g. Horn *et al*., 1994; Fonseca and Fleming, 1995; Srinivas and Deb, 1995; Weile *et al*., 1996; Cheng and Li, 1997; Zhou and Gen, 1999). Recently, Loughlin and Ranjithan (1997) presented a new multiobjective genetic algorithm called the neighbourhood constraint method (NCM) that uses a combination of a neighbourhood selection technique and location-dependent constraints. The authors demonstrate the NCM for complex, real-world problems, and their results show that NCM performs better than several other techniques including integer programming, single-objective GA and implementations of a Pareto and hybrid niched-Pareto

[2] In *GA-digest* Volume 6 number 32 (September 1992), the editor, Alan C. Schultz, lists various research using non-binary representations.

multiobjective GA. However, neither Loughlin and Ranjithan (1997) nor any other author working in the field of multiobjective GA have so far dealt with the measurement of efficiency for a specific object in a given data set.[3] This alternative way of identifying Pareto-optimal solutions will not be discussed here. Nevertheless, it must be admitted that such new research initiatives may provide very different alternative approaches for identifying comparative partners in the future.

8.3 Design of an EDEA

In the following, an evolutionary extension of DEA (EDEA) is proposed to solve typical DEA problems of efficiency measurement and of the selection of optimal benchmarking partners. This is a genetic algorithm (GA) approach. The challenging task is to make the optimization procedure usually employed in DEA more universally applicable especially for real-time decision support systems.

The EDEA model (EDEA) was programmed in C++ and is described in detail in the following sections. The reader interested in the source code may refer to the complete listing in Appendix B and is advised to contact the author in order to receive the latest version of the program. Any comments or advice on how to increase the efficiency of the program are appreciated.

8.3.1 Data structures and main program

The flow chart of the main EDEA program is displayed in Fig. 8.2. Starting from a population of randomly created initial chromosomes the genetic algorithm then manipulates the population of artificial chromosomes using the operations of reproduction, crossover, and mutation. Individuals are probabilistically selected to participate in these genetic operations based on their fitness.

The population is constructed as an array of individuals where each individual contains an array of parameters. These parameters include the chromosome itself (which are the values of the genes), a raw, a transferred and a scaled fitness value. The population is stored in the matrices *newpop[][]* and *oldpop[][]*, where the latter is a copy of the population generated in the last generation (except generation no. 1 where *oldpop[][]* is generated randomly). The dimensions of these matrices are spanned by *popsize*, the size of the population defined by the user, and the number of companies plus three columns for the various fitness values described above (*cases* + 3).

The main program is responsible for reading the DEA data from an input file, dynamically allocating memory according to the size of the model,

---

3 A list of references on Evolutionary Multiobjective Optimization is compiled by C.A. Coello and available on the Web via www.lania.mx/~ccoello/EMOO/EMOObib.html

initializing the EDEA and calling several subprograms for the calculation of temporary statistics, the actual optimization procedure and then reporting. The real data are read into the matrix *rec[][]* which is defined by the number of companies (*cases*) and the number of discretionary, non-discretionary input and output variables (*ianz*, *iuanz* and *oanz*).

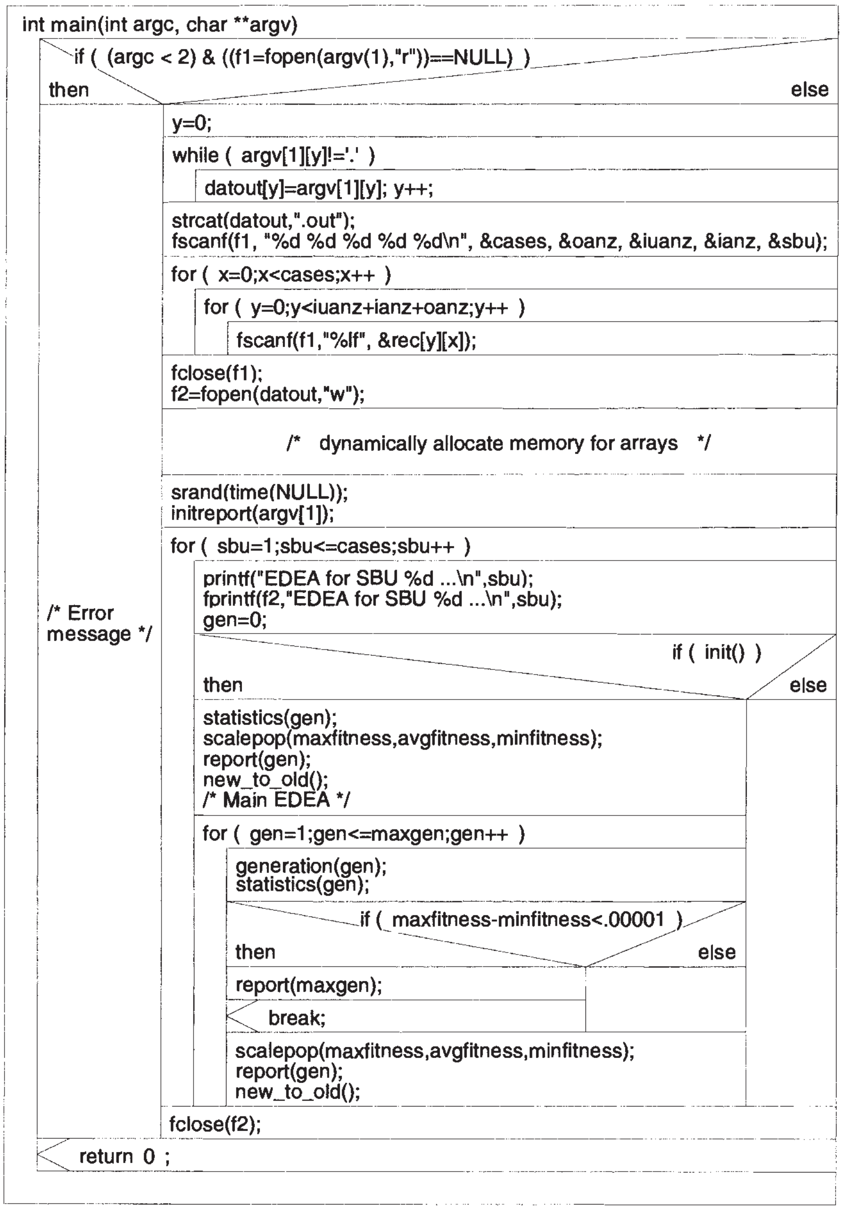


Fig. 8.2. EDEA main program.

In the main loop of the genetic algorithm, each individual in the population is evaluated for fitness and the population is tested for termination. If the termination criteria are not met, a new generation of the population is constructed by choosing one of the three genetic operations and performing that operation on individual(s) selected from the population based on fitness to create offspring individuals. Genetic algorithms are stochastic iterative processes that are not guaranteed to converge. Therefore, each run of the genetic algorithm requires specification of a termination criterion for deciding when to terminate a run and a method of result designation.

8.3.2 Constraint handling in EDEA

The proposed system adopts Michalewicz's GENOCOP approach, which provides a way of handling constraints that is both general and problem independent. It does not focus on the use of penalty functions as used by many other systems, rather it tries to keep all chromosomes within the constrained solution space by means of specially designed genetic operators. Furthermore, by reducing the type of constraints solely to inequalities, the search space becomes convex, which can be searched efficiently with some closed operators introduced by Michalewicz (1996). Since constant returns to scale DEA consists solely of inequalities in the constraints set, it is an easy task to apply this concept to DEA models. The variable returns to scale DEA, however, will not be discussed here.

The procedures required in EDEA for the handling of the constraints are *dea()*, *check_feasibilty()* and *calc_boundaries()*. The procedure *dea()*, which is illustrated in Fig. 8.3, is responsible for the calculation of the DEA efficiency score. The procedure requires the variables *oanz*, *iuanz* and *ianz*, which

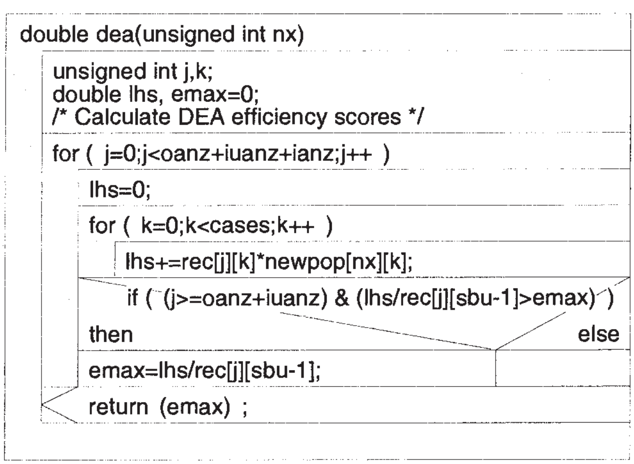


Fig. 8.3. Procedure for the calculation of the efficiency score.

represent the number of output, uncontrollable and controllable input variables of the DEA model. Note that this procedure does not need any penalty function.

The return value of procedure *dea()* is the efficiency score *emax*, which is calculated using the constant returns to scale DEA model, formulized in Equation 4.35 (p. 74).

The procedure *check_feasibility()* is an important procedure which is used in several instances of the program. This procedure, illustrated in Fig. 8.4, checks for any constraint violations for a specific individual in the population and returns a flag indicating the outcome of the evaluation.

The procedure calls for two subprograms. The first subprogram, *check_input()*, checks if the current gene values of the chromosome under evaluation violate any constraints which are defined by the uncontrollable variables in the model. The second subprogram, *check_output()*, searches for any output violation. Only if both subprograms have 'true' as a return value, is *check_feasibility()* also true.

The *calc_boundaries()* procedure is the third procedure that is problem specific in the general context of DEA modelling. The nature of the procedure is illustrated in Fig. 8.5. The procedure is required in many instances of EDEA. By means of the upper and lower boundaries the range is defined where a particular gene of a chromosome can be changed without violating any of the

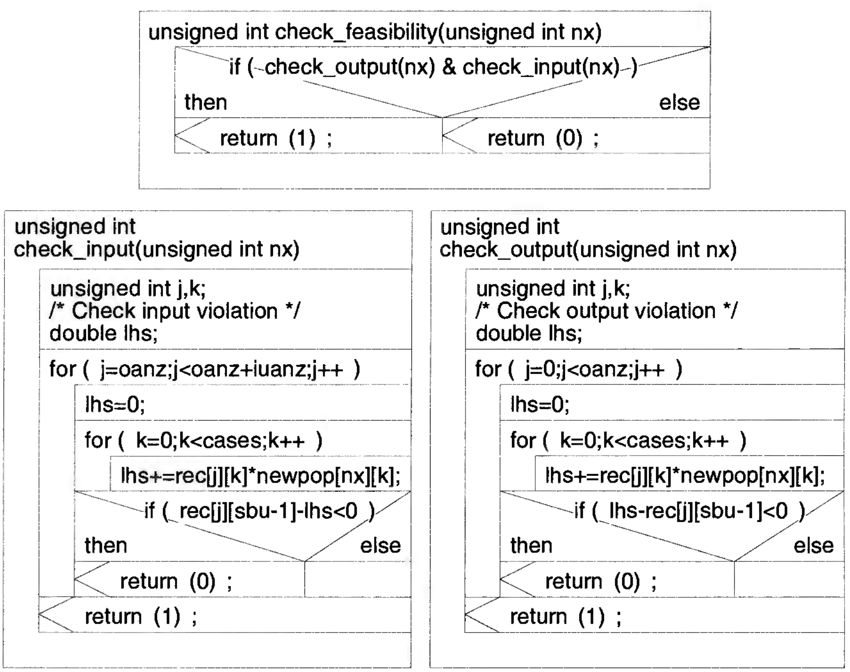


Fig. 8.4. Procedures checking for constraint violations.

constraints. The value of the *i*th gene of a feasible solution $s = \{v_1, \ldots, v_m\}$ is always in some dynamic range $[l, u]$, where the bounds l and u depend on the other vector's values $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m$, and the set of inequalities.

For the EDEA model the lower bounds are limited by the output constraints and the upper bounds depend on the non-discretionary input constraints defined by the model.

The lower bounds are easily calculated by

$$l^s_{(i)} = \max_j \frac{y^s_{ij} - \sum_{i=1}^{cases} L^s_i y^s_{ij} + y^s_{oj}}{y^s_{ij}} \qquad j = 1, \ldots, oanz;\ i = 1, \ldots, cases;\ \text{and } l^s_{(i)} \geq 0 \quad (8.1)$$

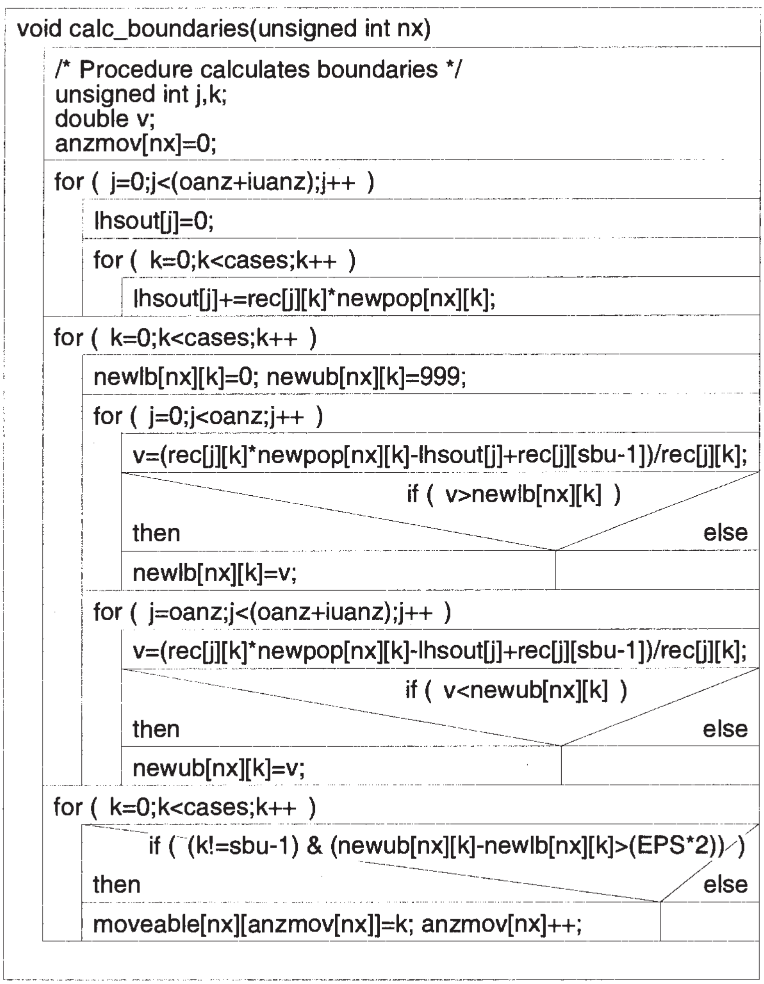


Fig. 8.5. Procedure for calculating the boundaries of the solution space.

and the upper bounds are defined by

$$u^s_{(i)} = \min_r \frac{z^s_{ir} - \sum_{i=1}^{cases} L^s_i z^s_{ir} + z^s_{or}}{z^s_{ir}} \quad r = 1, \ldots, iuanz;\ i = 1, \ldots, cases;\ \text{and}\ u^s_{(i)} \leq 1 \quad (8.2)$$

Calc_boundaries() is the implementation of Equations 8.1 and 8.2. It also maintains a list of all moveable genes of all chromosomes, which is required later by some of the genetic operators. A gene is defined as being moveable when its upper and lower boundaries are unequal, hence leaving some possibility for a genetic change. The number of moveable genes for a chromosome *i* is stored in *anzmov[i]*. For each moveable gene *k* the location within *i* is stored in the matrix *moveable[i][k]*. Note from the program that $\sum_{i=1}^{cases} L^s_i y^s_{ij}$ is the left-hand-side sum of each output constraint, which are temporarily stored in the array *lhsout[]*.

8.3.3 Initialization

For the first generation, the EDEA main program selects potential solutions randomly from the space of all feasible solutions. To guarantee a feasible solution a feasibility check is performed for each randomly generated chromosome. If the chromosome survives the feasibility check it is moved to the initial mating pool and the procedure is started again until the number of chromosomes reach the population-size, *popsize*, defined by the user. If the chromosome is an infeasible solution then it is rejected and a new chromosome is generated randomly. This search of an initial set of potential solutions is repeated until a user-defined value is reached, thus terminating the initialization procedure and returning a false value. This situation usually occurs in a simplex-based program for efficient companies where no (Andersen and Petersen-) 'superefficiency score' can be calculated (= infeasible solution).

The procedure *init()*, displayed in Fig. 8.6, can be slow when the solution space is fragmented and heavily constrained, however, experiments so far have indicated that this is usually not the case in real-world DEA problems.

In the procedure *calc_boundaries()* the boundaries of the present solution are calculated and stored into the matrix *new_lb[][]* for the lower bounds, and *new_ub[][]* for the upper bounds of each individual gene of all chromosomes in the population (in analogy, *newpop[][]* and *oldpop[][]*, *old_lb[][]* and *old_ub[][]* are copies of the last generation's upper and lower bounds).

The initialization starts with the generation of a random population and continues with the calculation of upper and lower bounds. The next step is the calculation of the initial population statistics, and the printing of a special initial report using the procedures *statistics()* and *initreport()*. In *statistics()*, indicators that monitor important attributes of the current population are calculated and stored in global variables which are required for fitness scaling, reporting and for deciding on program termination. Furthermore, the procedure is responsible for the transformation of the efficiency scores evaluated in

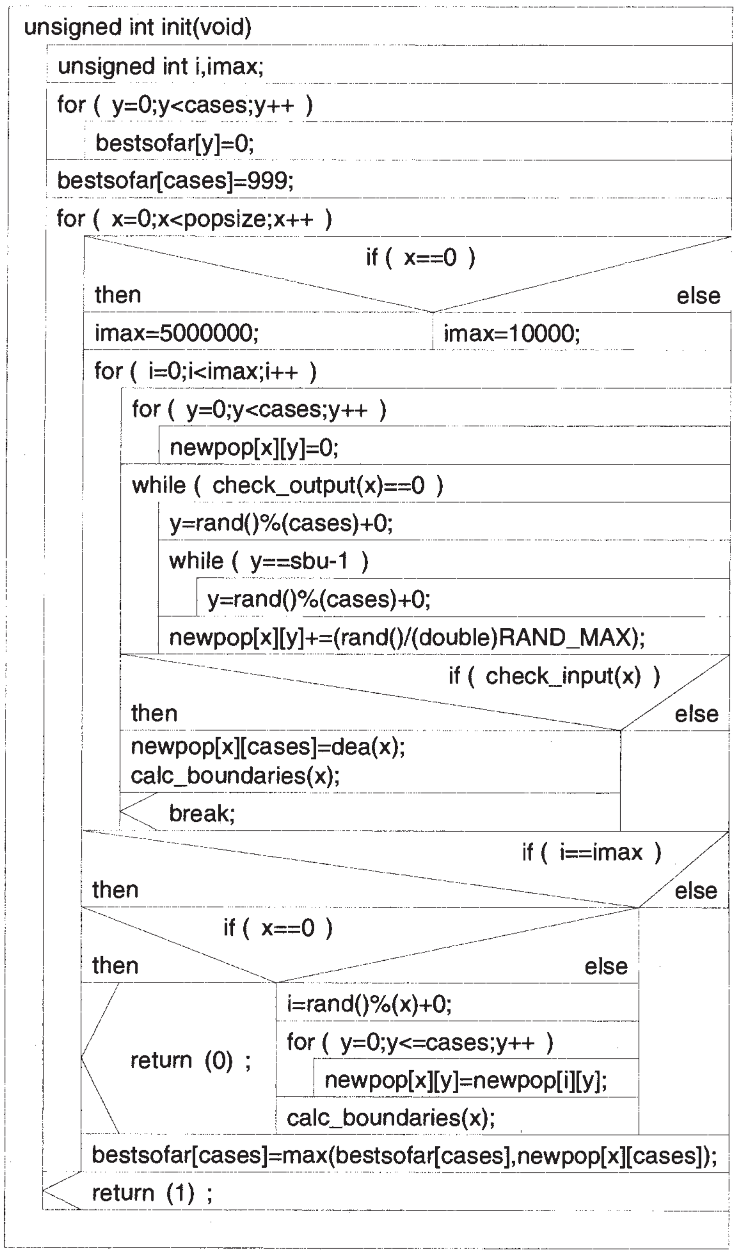


Fig. 8.6. Procedure generating the initial generation.

dea() into fitness scores, which are used by the GA. This is necessary as the objective in an input-oriented DEA is a minimization of an efficiency function $e(x)$ rather than the maximization of a fitness function $f(x)$. As a result, it is necessary to convert the underlying objective function into a fitness function form through a mapping procedure.

In GA the fitness function must be guaranteed to be non-negative in all instances. Thus, simply multiplying the cost function by −1 does not work. In EDEA the following efficiency-to-fitness transformation is used:

$$f(x) = e_{\max}(x) * c - e(x) \qquad \text{where const } c > 1 \tag{8.3}$$

Here, $e_{\max}(x)$ is the largest efficiency score observed in the current population, and the constant c is included to guarantee a positive value and that there is at least a minimum chance for each chromosome to be chosen by the subsequent generation.

The indicators and their variable names generated in *statistics()* are:

- the maximum scaled fitness of the current population (*maxfitness*);
- the reference number of the best chromosome in the current population (*bestchrom*);
- the minimum scaled fitness of the current population (*minfitness*);
- the average scaled fitness of the current population (*avgfitness*); and
- the total sum of scaled fitness of the current population (*sumfit*).

This version of *statistics()* is something of a minimally acceptable solution. Many other interesting population statistics could be tracked. For example, the best chromosome so far could be stored for future reference. Population standard deviation or even population histograms might also be of interest in doing more detailed run postmortems. The separation of statistical functions in the routine statistics permits the easy addition of any or all of these computations. Therefore, the test will not dwell further on the code of *statistics()* and the two reporting procedures. The interested reader should refer to Appendix B, which contains a complete copy of the EDEA code.

At long last, with all the necessary preliminaries complete, the program can start with the main loop contained within the do-until construct. There are two termination criteria which have been built into EDEA. First, the program terminates when a maximum number of generations (*maxgen*), which is defined by the user, is reached. Second, the program will end when the difference between the average and the minimum fitness of the current population falls below a certain value (e.g. 0.0001), which is a strong indication that the population reached uniformity (converged to an optima). In rapid succession, in the main loop the generation counter (*gen*) is incremented, a new generation is created, new generation statistics are calculated, the scaling of the new fitness values is performed, and a generation report is printed. If the termination check fails then the program continues by swapping the current population variables into the old population variables in procedure *new_to_old()* (see Appendix B for details of the system code). All this continues until one of the two termination criteria is fulfilled, thereby forcing the machinery to halt.

The procedure *report()* presents the full population report, including chromosomes, various fitness, and upper and lower bound values. An output listing produced by *report()* is presented in Appendix B. Again, many tabular or graphic reporting options may be useful for presenting the EDEA results. The simple report procedure implemented here is a good tool because it permits the investigation of the moveable genes and the inspection of the effectiveness of the genetic operators.

8.3.4 The GA operators

The main iterative part of the program where the GA operators' reproduction, crossover and mutation are invoked, is placed in the procedure *generation()*, which is illustrated in Fig. 8.7. The main loop starts by generating a random number which decides on the GA operator which will be processed next. Each operator is assigned with a different probability which are stored in variables and belong to the set of parameters the user has to define before program start. Depending on the type of operation, the program will generate one or two offspring, which are added to the new population. Once the number of chromosomes has reached *popsize*, the program is terminated and control is returned to the main program.[4]

EDEA consists of eight GA operators, which are listed below and will now be discussed in detail. The three crossover and four mutation operators suggested by Michalewicz's GENOCOP system have been complemented by an additional mutation operator suggested by the author which ought to be specially useful for DEA problems. The operators implemented in EDEA with their probability parameters are:

1. The reproduction operator (*default*);
2. The whole arithmetical crossover (*pm_wa*);
3. The simple crossover (*pm_sc*);
4. The single arithmetical crossover (*pm_ac*);
5. The uniform mutation (*pm_um*);
6. The boundary mutation (*pm_bm*);
7. The non-uniform mutation (*pm_nm*); and
8. The conditional mutation (*pm_cm*) suggested by the author.

Reproduction

In the reproduction operation, an individual is probabilistically selected from the population based on its fitness (with reselection allowed) and then

[4] To avoid an array overflow during a crossover operation when only one chromosome is missing, the populations' arrays are defined accordingly (*popsize* +1).

the individual is copied, without change, into the next generation of the population.

```
void generation(unsigned int gen)
    /* Generates a new population through select, crossover, and mutation */
    /* GA operators return increment in population size */
    unsigned int popincr,j,k;
    double rnumber;
    x=0;
    while ( x<popsize )
        j=select();
        rnumber=rand()/(double)RAND_MAX;
        switch ( rnumber )
            case > pm_wa+pm_sa+pm_sc+pm_cm+pm_um+pm_bm+pm_nm :
                popincr=reproduction(x,j);
            case > pm_wa+pm_sa+pm_sc+pm_cm+pm_um+pm_bm :
                popincr=mutation_nm(x,j);
            case > pm_wa+pm_sa+pm_sc+pm_cm+pm_um :
                popincr=mutation_bm(x,j);
            case > pm_wa+pm_sa+pm_sc+pm_cm :
                popincr=mutation_um(x,j);
            case > pm_wa+pm_sa+pm_sc :
                popincr=mutation_cm(x,j);
            default:
                k=select();
                while ( k==j )
                    k=select();
                switch ( rnumber )
                    case > pm_wa+pm_sa :
                        popincr=crossover_sc(x,j,k);
                    case > pm_wa :
                        popincr=crossover_sa(x,j,k);
                    default:
                        popincr=crossover_wa(x,j,k);
        x+=popincr;
    for ( j=0;j<popsize;j++ )
        newpop[j][cases]=dea(j);
        calc_boundaries(j);
```

Fig. 8.7. The main loop in EDEA.

Not only the reproduction itself, but also all other operators require the selection of one (in the case of a mutation) or two (in the case of a crossover) chromosome(s). Many selection procedures are currently in use, the two most popular ones are the 'roulette wheel selection' and the 'tournament selection'. The selection process implemented in EDEA is the roulette wheel selection which is shown in Fig. 8.8.

The roulette wheel procedure is Holland's original fitness-proportionate selection. In this procedure individuals are selected with a probability proportional to their relative fitness which is like a roulette wheel with slot sizes proportional to the fitness. This ensures that the expected number of times an individual is chosen is approximately proportional to its relative performance in the population. Thus, high-fitness individuals stand a better chance of reproducing their genes, whereas low-fitness ones are more likely to disappear.

The procedure is defined as follows. Because the population size is usually kept fixed in a GA, the cumulative probability for all strings in the population is one. The probability for selecting the ith string is therefore defined as

$$\frac{f_i}{\sum_{j=1}^{N} f_j} \tag{8.4}$$

where N is the population size. The roulette-wheel is spun N times, each time selecting one individual by the roulette-wheel pointer. The roulette-wheel mechanism is expected to move f_i/f_{avg} copies of the ith chromosome from the old to the new population (where f_{avg} denotes the average fitness in the old population).

This kind of selection incorporates quite a natural approach, as fitness in natural systems is often interpreted as the ability to survive and multiply.

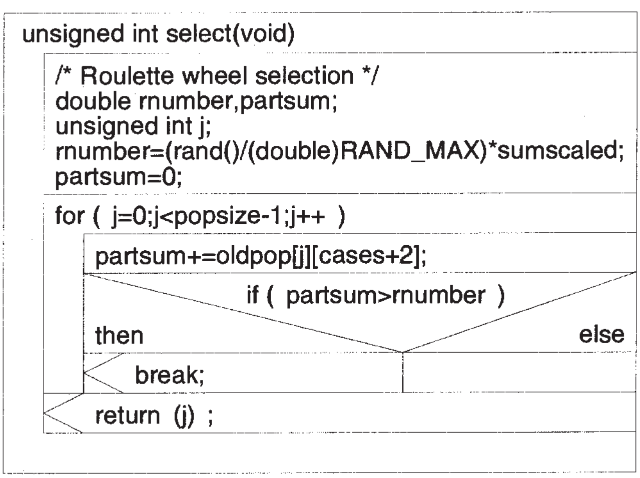


Fig. 8.8. Roulette wheel selection in EDEA.

Another important aspect of probabilistic selection is that every individual, however poor, has some probability of selection.

In principle, the roulette wheel selection is a powerful and fast algorithm tool. However, problems with premature convergence and slow finishing have to be avoided (Beasley *et al.*, 1993). In the beginning of a GA run it frequently happens that a chromosome receives a relatively high fitness (compared to other chromosomes in the population) even though it is neither optimal nor near optimal. The proportional selection of the roulette wheel procedure spreads this chromosome into the population very quickly, thus, it could easily happen that after a few generations the population consists entirely of this chromosome. This phenomenon when the GA gets stuck in an unfavourable situation is called premature convergence.

The other problem with proportional selection, slow finishing, refers to the fine tuning at the end of the search. In a case where the average fitness is very close to the highest possible fitness, a string encoding the optimal solution is very rarely expected. Considering the disruptive effect of the other genetic operators, it is highly unlikely that this optimal solution will take over the whole population.

Premature convergence and slow finishing can be avoided by using scaled fitness values instead of the original fitness values in the selection process. The most popular kind of scaling is the linear scaling, which was suggested by Goldberg (1989: 76). Goldberg proposes a linear relationship between the raw fitness value f and the scaled fitness value f' as follows:

$$f' = af + b \tag{8.5}$$

The parameters a and b are calculated for each generation in order to satisfy the following two equalities:

$$f'_{avg} = f_{avg} \text{ and } f'_{max} = f_{multiple} f_{avg} \tag{8.6}$$

where f_{avg} is the average, f_{max} is the maximal fitness in the population and $f_{multiple}$ is a parameter which determines the selection pressure. The relationship of raw and scaled fitness is illustrated in Fig. 8.9. Typical values of $f_{multiple}$ are between 1.2 and 2. If $f_{multiple}$ has a value of 2, then the string with the highest fitness in the population expects 2 offspring regardless of the difference between maximal and average payoff.

However, $f_{multiple}$ can stretch the raw fitness significantly, usually at the end of a run, and causes another problem as some of the extraordinarily low fitness values may go negative after scaling (see Fig. 8.10). One of the solutions for this problem is to map the minimum raw fitness f_{min} to a scaled fitness $f'_{min} = 0$ instead of $f_{multiple}$.

The linear scaling procedure as illustrated in Fig. 8.11 takes the average, maximum and minimum raw fitness values and calculates linear scaling coefficients a and b based on the logic described by Equations 8.5 and 8.6. If it is possible to scale to the desired multiple, $f_{multiple}$, then that is the computation performed. Otherwise, scaling is performed by pivoting about the average value and stretching the fitness until the minimum value maps to zero. In the following loop all the individuals' raw fitness values are adjusted and stored in a separate column of *newpop[][]*, next to the row fitness values. The sum of

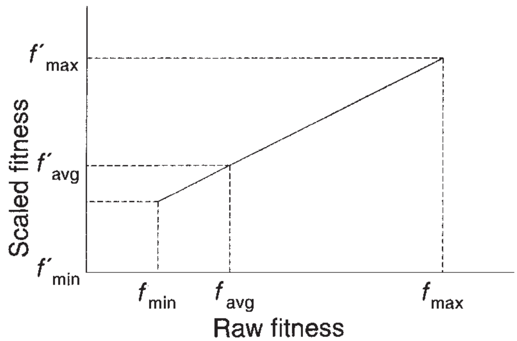


Fig. 8.9. Linear scaling (Goldberg, 1989).

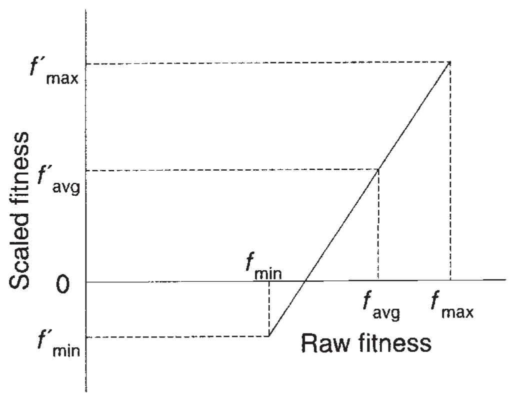


Fig. 8.10. Negative fitness violation in the scaling procedure suggested by Goldberg (1989).

```
void scalepop(double umax, double uavg, double umin)
    unsigned int j;
    double a,b;
    const double fmultiple=4;
    /* Calculate scaling coefficients for linear scaling (Goldberg 1989) */
    if ( umax==umin )
    then
        a=1;
        b=0;
    else
        if ( umin>(fmultiple*uavg-umax)/(fmultiple-1) )
        then
            a=(fmultiple-1)*uavg/(umax-uavg);
            b=uavg*(umax-fmultiple*uavg)/(umax-uavg);
        else
            a=uavg/(uavg-umin);
            b=-umin*uavg/(uavg-umin);
    sumscaled=0;
    for ( j=0;j<popsize;j++ )
        newpop[j][cases+2]=max(a*newpop[j][cases+1]+b,0);
        sumscaled+=newpop[j][cases+2];
```

Fig. 8.11. The scaling procedure.

the scaled fitness values *sumscaled* is recalculated. This simple scaling helps to prevent the early domination of extraordinary individuals, and later on it encourages a healthy competition among near equals.

Selection alone cannot introduce any new individuals into the population, i.e. it cannot find new points in the search space. These are generated by genetically inspired operators, of which the most well known are crossover and mutation.

Crossover operators

The genetic operation of crossover allows new individuals, which represent new points in the search space, to be created and tested. Inspired by the example of nature, crossover is intended to join the genetic material of two chromosomes, called parents, with a high fitness in order to produce even better individuals, referred to as children (or offspring). In general, the crossover operator tends to enable the evolutionary process to move toward promising regions of the search space. A theoretical foundation why the crossover operator increases the performance of a GA is given by Holland's Schema Theorem and the building block hypothesis (Holland, 1975).

As in the reproduction phase, the selection for crossover is done in such a way that the better an individual's fitness, the more likely it is to be selected. There are several crossover procedures suggested by various authors, but all approaches have one thing in common; each offspring contains some genetic material from each of its parents.

In the simplest binary coded form, the one-point crossover, two substrings are exchanged after a randomly selected crossover point. This is performed by randomly choosing a crossing site along the string and by exchanging all bits on one of the two sides of the crossing point as shown:

$$00|000 \Rightarrow 00|111$$
$$11|111 \Rightarrow 11|000$$

Since the knowledge of an appropriate crossing point is usually not known, the crossing point has to be selected randomly. Therefore, it is not sure if the children chromosomes produced will be better than each of the parent's fitness. However, this is not a major problem in GAs as only good chromosomes created by crossover will survive in the reproduction phase of the entire procedure.

Apart from one-point crossover, a number of other crossover techniques were introduced in the GA literature. Some researchers use two-point or even multi-point crossover, where a certain number of crossover points are chosen, and the genetic material is swapped between every two of these points (Kinnebrock, 1994: 76). In the beginning of GA research many studies were performed to find an optimal crossover operator. However, most of these studies which recommend a specific crossover operator are very problem dependent and cannot be generalized enough to be used for all problems (Deb, 1996).

EDEA uses three crossover operators suggested by Michalewicz (1996) which he integrated and tested in the GENOCOP system. All three operators have in common the attribute that they can be used in GAs with floating point representation and that they guarantee offspring which stay within the feasible area of the problem space. The three operators, which are the whole arithmetical crossover, the simple crossover and the single arithmetical crossover, are explained briefly in the following.

WHOLE ARITHMETICAL CROSSOVER The whole arithmetical crossover is defined as a linear combination of two chromosomes. Hence, in a case where the chromosomes ${s^t}_v$ and ${s^t}_w$ are to be crossed, the resulting offspring are

$$s_v^{t+1} = a \cdot s_w^t + (1-a) \cdot s_v^t \text{ and } s_w^{t+1} = a \cdot s_v^t + (1-a) \cdot s_w^t \tag{8.7}$$

The arithmetical crossover operator uses a simple static system parameter $pm\_a \in [0..1]$, as it always guarantees that the result stays within the feasible region of the solution space. The procedure is illustrated in Fig. 8.12.

Other values of *pm_a* may be considered so that, instead of the new value being interpolated in that range, the new value could be obtained by extrapolating the two values. Eshelman and Schaffer (1993), for instance, have introduced a BLX-0.5 crossover operator with *pm_a* varying between [–0.5..1.5].

SIMPLE CROSSOVER In the simple crossover procedure, the chromosomes ${s_v}^t = \{v_1, \ldots, v_m\}$ and ${s_w}^t = \{w_1, \ldots, w_m\}$ are crossed after the *k*th position and the resulting offspring are ${s_v}^{t+1} = \{v_1, \ldots, v_k, w_{k+1}, \ldots, w_m\}$ and ${s_w}^{t+1} = \{w_1, \ldots, w_k, v_{k+1}, \ldots, v_m\}$.

However, such an operator may produce offspring outside the convex solution space S. To avoid this problem, Michalewicz suggests using the property of convex spaces where $a \in [0..1]$ exists such that

$$\begin{aligned} s_v^{t+1} &= \{v_1, \ldots, v_k, w_{k+1} \cdot a + v_{k+1} \cdot (1-a), \ldots, w_m \cdot a + v_m \cdot (1-a)\} \in S \\ s_w^{t+1} &= \{w_1, \ldots, w_k, v_{k+1} \cdot a + w_{k+1} \cdot (1-a), \ldots, v_m \cdot a + w_m \cdot (1-a)\} \in S \end{aligned} \tag{8.8}$$

When $a = 1$ the two chromosomes will completely exchange their genes starting from the *k*th position; when $a = 0$ no changes are made. In order to achieve the highest possible information exchange, EDEA performs a step-wise search

```
unsigned int crossover_wa(unsigned int x, unsigned int mate1, unsigned int mate2)
    /* Whole arithmetical crossover (Michalewicz 1996) */
    unsigned int j;
    wac++;
    for ( j=0;j<cases;j++ )
        newpop[x][j]=oldpop[mate1][j]*pm_a+oldpop[mate2][j]*(1-pm_a);
        newpop[x+1][j]=oldpop[mate2][j]*pm_a+oldpop[mate1][j]*(1-pm_a);
    return (2) ;
```

Fig. 8.12. Whole arithmetical crossover.

starting with $a = 1$ and, in case of infeasibility, by reducing it by a user-defined constant number from the range [0..1]. This is an extension to the original approach in the GENOCOP system where Michalewicz suggested an ordinary binary search tool.

A possible coding of the simple crossover procedure is illustrated in Fig. 8.13. First, the crossover point is generated randomly and stored in the variable *jcross*. Second, the genes of the parents' chromosomes are copied to the new chromosome until the *jcross*-th gene has been reached. Inside the loop of *crossover_sc()* the program looks for a high value for a. Here, the offspring are defined by following Equation 8.8 and checked for feasibility unless a feasible solution has been found or $a = 0$. In the procedure displayed in Fig. 8.13 the variable a is decreased by a constant value 0.2.

SINGLE ARITHMETICAL CROSSOVER In the single arithmetical crossover the chromosomes $s_v^t = \{v_1, \ldots, v_m\}$ and $s_w^t = \{w_1, \ldots, w_m\}$ are crossed at a particular position k by exchanging only a single gene. The resulting offspring are $s_v^{t+1} = \{v_1, \ldots, v'_k, \ldots, v_m\}$ and $s_w^{t+1} = \{w_1, \ldots, w'_k, \ldots, w_m\}$, where $k \in [1..m]$, $v'_k = a \cdot w_k + (1 - a) \cdot v_k$, and $w'_k = a \cdot v_k + (1 - a) \cdot w_k$.

The parameter a is randomly selected within a given range which guarantees that the outcome stays within the feasible region of the solution space. The range from which a is randomly selected is defined by:

$$a \in \begin{cases} [\max(\alpha,\beta), \min(\gamma,\delta)] & \text{if } v_k > w_k \\ [0,0], & \text{if } v_k = w_k \\ [\max(\gamma,\delta), \min(\alpha,\beta)] & \text{if } v_k < w_k \end{cases} \tag{8.9}$$

```
unsigned int crossover_sc(unsigned int x, unsigned int mate1, unsigned int mate2)
  /* Simple crossover (Michalewicz 1996) */
  unsigned int jcross,j;
  double a=1;
  scc++;
  jcross=rand()%(cases)+0;
  for ( j=0;j<jcross;j++ )
     newpop[x][j]=oldpop[mate1][j]; newpop[x+1][j]=oldpop[mate2][j];
  if ( jcross<cases )
  then
  while ( a>=0 )
     for ( j=jcross;j<cases;j++ )
        newpop[x][j]=oldpop[mate2][j]*a+oldpop[mate1][j]*(1-a);
        newpop[x+1][j]=oldpop[mate1][j]*a+oldpop[mate2][j]*(1-a);
     if ( check_feasibility(x)+check_feasibility(x+1)==2 )
     then
        break;
     else
        a-=.2;
  else
  return (2) ;
```

Fig. 8.13. The simple crossover procedure.

where

$$\alpha=\frac{\left(l_{(k)}^{sw}-w_k\right)}{\left(v_k-w_k\right)},\ \beta=\frac{\left(u_{(k)}^{sv}-v_k\right)}{\left(w_k-v_k\right)},\ \gamma=\frac{\left(l_{(k)}^{sv}-v_k\right)}{\left(w_k-v_k\right)},\ \delta=\frac{\left(u_{(k)}^{sw}-w_k\right)}{\left(v_k-w_k\right)} \tag{8.10}$$

The coding which implements the single arithmetical crossover in EDEA is illustrated in Fig. 8.14. The value of *a*, denoted with *cincr* in the

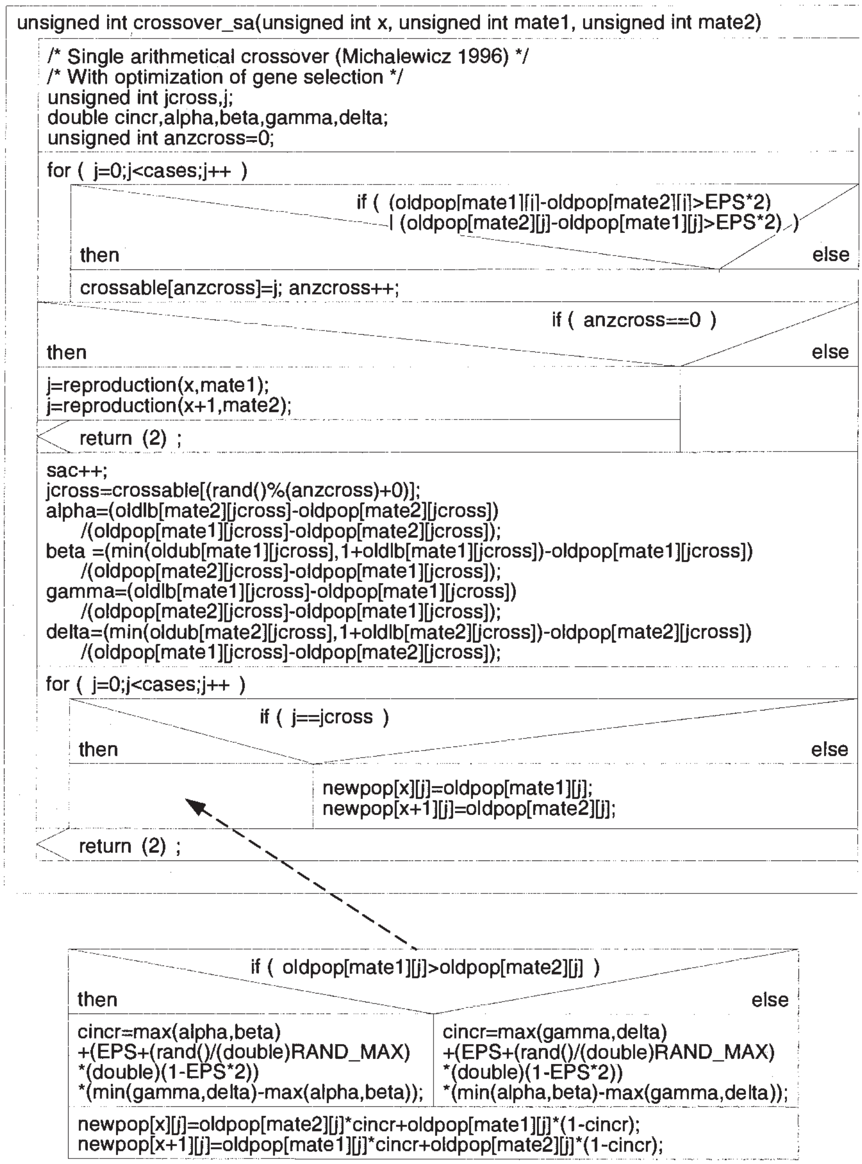


Fig. 8.14. Single arithmetical crossover.

program, must be determined separately for each procedure call and for each gene.

The significance of the genetic change caused by this operator varies with the difference between the two selected genes. If the difference is small the genetic change will be less significant and vice versa. If both genes show the same value, the single arithmetical crossover will be completely ineffective. To avoid this unfavourable situation the program selects only genes which show different weights.

At this point one note is appropriate concerning the problem of floating point numbers which may lose precision with commonly used C++ compilers. Towards the end of the EDEA run when the movable range of parameter values gets tighter and the values approach the upper and lower bounds, the α and γ variables reach very large values. In this case floating point calculations in C++ can lead to unexpected results. The reason is because floating point decimal values generally do not have an exact binary representation in C++. This is a side effect of how the CPU represents floating point data. This especially can become a problem in a highly iterative procedure like a GA and where non-negativity is required in several instances of the program. There is nothing much to do about this problem besides building in user-defined tolerances for calculations involving large numbers. For this reason, *eps* (epsilon), a very small constant, was added or subtracted in the calculations of the upper and lower bounds respectively, to avoid anomalies that could occur in the calculation of the single arithmetical crossover as well as in all mutation procedures.

Mutation operators

Whereas the reproduction operator reduces the diversity in the population, the mutation operator increases it again. Furthermore, for fine local improvement of a solution at the very end of a GA run, mutation is usually very useful.

Similar to the crossover operator, the mutation procedure begins by selecting an individual from the population based on its fitness. A point along the string is selected at random and the character (or value in a floating point presentation) at that point is randomly changed into one of the other characters from the alphabet (values from the range).

Usually, mutation in GA is used with some small probability. By randomly sampling new points in the search space, a high mutation rate would reduce the danger of premature convergence to local optima, however, caution must be taken as this may transform a GA to a pure random search algorithm, which is of course not the intention of the algorithm.

Michalewicz suggested three different types of mutation operators which he integrated in the GENOCOP system. These are uniform mutation, boundary mutation and non-uniform mutation. Due to some weaknesses in the uniform mutation and peculiarities of DEA optimization problems the author suggests a new mutation operator, referred to as ‘conditional mutation’.

UNIFORM MUTATION When uniform mutation is selected by the program, a gene is mutated within the feasible range. Formally spoken, if $s_v^t = \{v_1, \ldots, v_m\}$ is a chromosome and the kth component is the selected gene, the result is the chromosome $s_v^{t+1} = \{v_1, \ldots, v'_k, \ldots, v_m\}$, where v'_k is a random value following a uniform probability distribution from the range $\left[l_{(k)}^{s_v^t}, u_{(k)}^{s_v^t} \right]$. The coding of this simple uniform mutation procedure is illustrated by Fig. 8.15.

In order to improve the effectiveness of this operator only moveable genes are selected for a uniform mutation. If none of the genes of the selected chromosome are moveable, then a simple reproduction is performed.

BOUNDARY MUTATION Boundary mutation is a variation of uniform mutation with v'_k being either $l_{(k)}^{s_v^t}$ or $u_{(k)}^{s_v^t}$ with equal probability. The boundary mutation is an important mutation operator as it searches for potential solutions in very extreme locations of the solution space. This is a meaningful contrast to the whole arithmetical crossover which tends to uniform all chromosomes in a population. Also in the boundary mutation only moveable genes are selected for the operation. The program coding of this procedure is very simple and illustrated in Fig. 8.16.

NON-UNIFORM MUTATION In non-uniform mutation the action of the mutation process depends on the age of the population. It is an operator responsible for the fine tuning capabilities of the system. The non-uniform mutation, defined by Michalewicz (1996), is applied in EDEA on a chromosome $s_v^t =$

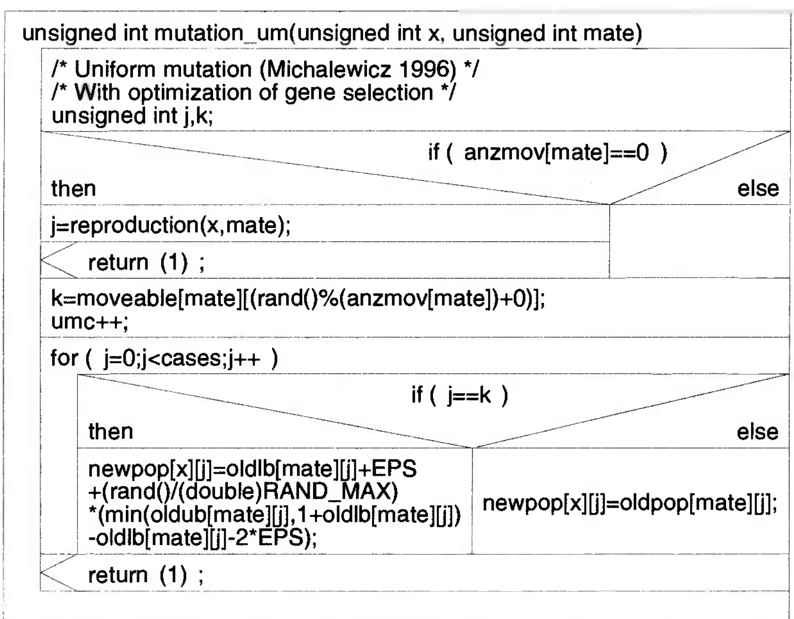


Fig. 8.15. Uniform mutation.

$\{v_1, \ldots, v_m\}$ by selecting the kth element and mutating it into $s_v{}^t = \{v_1, \ldots, v'_k, \ldots, v_m\}$, with $k \in \{1, \ldots, cases\}$, and

$$v'_k = \begin{cases} v_k + s\left(t, u_{(k)}^{s_v^t} - v_k\right) & \text{if a random digit} = 0 \\ v_k + s\left(t, v_k - l_{(k)}^{s_v^t}\right) & \text{if a random digit} = 1 \end{cases} \tag{8.11}$$

The function $s(t,y)$ is any simulated annealing algorithm which returns a value in the range $[0..y]$ such that the probability of $s(t,y)$ being close to 0 increases as t increases. This property causes this operator to search uniformly initially (when t is small), and very locally at later stages. The simulated annealing algorithm used here is defined by

$$s(t,y) = y \cdot \left(1 - r^{\left(1 - \frac{t}{T}\right)^b}\right) \tag{8.12}$$

where r is a random number from $[0..1]$, T is the maximal generation number, and b is a system parameter determining the degree of non-uniformity.

The coding of the non-uniform mutation operator is illustrated in Fig. 8.17. This procedure, *mutation_nm()*, calls the function *cooling()* (see Appendix B) which implements the simulated annealing algorithm defined in Equation 8.12. The parameter b, which is denoted with *pm_b* in EDEA, must be defined before program start; however, a value of 2 is recommended. Also the non-uniform mutation operation contains some code which guarantees that only moveable genes are selected for this procedure.

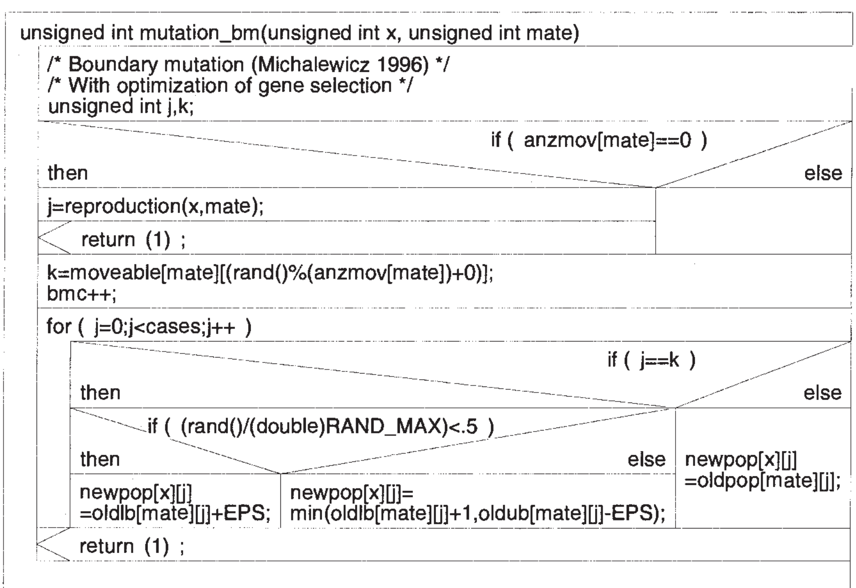


Fig. 8.16. Boundary mutation.

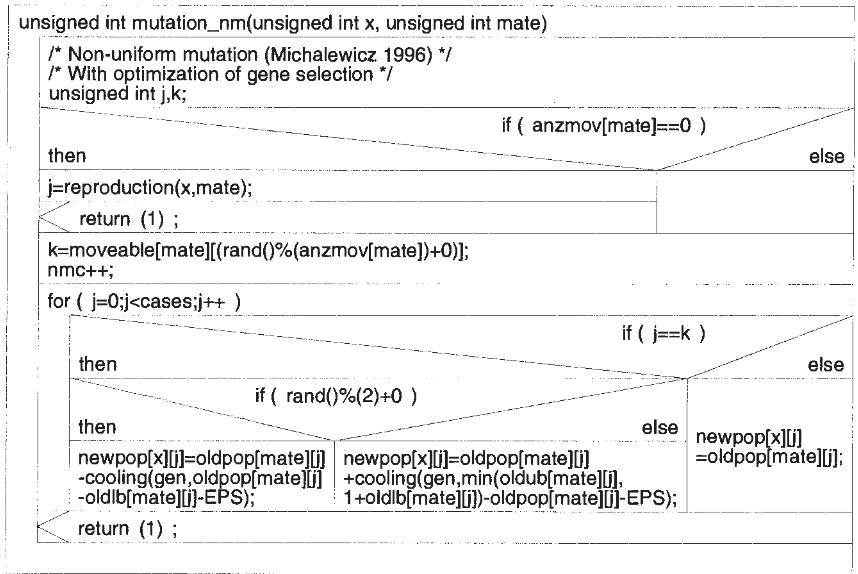


Fig. 8.17. Non-uniform mutation.

CONDITIONAL MUTATION Mutations in constraint genetic algorithms cannot be embedded in crossover or reproduction operations as suggested by many authors for non-constraint, ordinary optimization problems (e.g. Goldberg, 1989). In systems where a mutation operator is embedded in a crossover or reproduction operator, several genes of a single chromosome have at least a certain chance to be mutated simultaneously. However, this is not possible in a constraint environment where genetic operators are dynamic, i.e. a value of a chromosome component depends on the remaining values of the other genes. Therefore, the type of mutation performed on a gene has to be decided individually and multiple operations on one and the same chromosome are not possible. This is problematic in an environment where there are multiple solutions near optima because, once trapped in a local optimum, the GA may have severe difficulties in leaving this area again.

In the GENOCOP system it was not really necessary to emphasize this problem as five out of six test functions of a transportation problem introduced by Michalewicz (1996) consisted of a unimodal format.

However, for the present DEA application it was demonstrated in Section 8.1.1 that there are potentially hundreds or thousands of instances where a DEA program may get trapped when all weights are located on the boundary of the solution space. Unfortunately, the desperately searched global optimum is also located on the boundary, so that it is also preferable for the algorithm to search in these extreme areas.

None of the suggested mutation operators suggested by Michalewicz are designed to leave local optima easily. In fact, the only way to leave one of the suboptimal solutions presented in Section 8.1.1 is for several disadvantageous operations to be carried out consecutively on the same individual. However, the probability that this individual with a (temporarily) very bad fitness will survive in the next generation is very small, especially during the end of the run (when the variance in the fitness values is small).

To overcome these weaknesses the author proposes an additional mutation operator which allows the simultaneous variation of two genes. The procedure starts by selecting two genes randomly for information exchange. If $s_v{}^t = \{v_1, \ldots, v_m\}$ is a chromosome and the kth and lth components are the selected genes, the offspring is $s_v{}^{t+1} = \{v_1, \ldots, v'_k, \ldots, v'_l, \ldots, v_m\}$, where

$$v'_l = \begin{cases} v_l, \\ v_l - r \cdot a \cdot v_l, \end{cases} \quad \text{and} \quad v'_k = \left.\begin{cases} v_k, & \text{if } v_k = v_l \\ v_l + r \cdot a \cdot v_l, & else \end{cases}\right\} \in S \tag{8.13}$$

and r is a random number from [0..1]. If v'_k and v'_l are simply modified, then it would be very likely that this would produce an offspring outside the convex solution space S. To avoid this problem, the property of the convex solution space can be used, in analogy to the simple crossover operator (see p. 174), where there exists $a \in [0..1]$ so that Equation 8.13 produces a feasible solution.

When $a = 1$, the two genes will exchange information in the way that a proportion of the v_l weight is transferred to v_k; when $a = 0$, no changes are made. In order to achieve the highest possible information exchange EDEA performs a stepwise search starting with $a = 1$ and reducing stepwise by a user-defined constant until a feasible solution is found or $a = 0$. An illustration of how the conditional mutation is coded in EDEA is presented by Fig. 8.18. In order to raise the effectiveness of this operator, the program selects gene l among $v_m > 0$.

8.4 Testing EDEA

It is difficult to generalize experimental results when trying to make some global claim about a particular technique. However, here it is possible to demonstrate the utility of a new method by comparing it against other well-established techniques. As the simplex algorithm has been the preferred technique for DEA in the past it was selected for systematic comparison with the GA counterfeit.

In this section the EDEA algorithm is compared with the ordinary linear programming solution (LP-DEA) which was produced with EMS and DEAP, the special software packages for running DEAs used previously in Chapter 7. Comparing LP-DEA and EDEA results is a very difficult test for EDEA as the simplex algorithm in ordinary DEA programs, with the multistage option for slack optimization, guarantees the finding of a global optimum.

Hypotheses regarding the proposed system are embodied in the following five working assumptions.

1. EDEA finds similar solutions compared to the global optimum found by the simplex-DEA.
2. EDEA produces high-quality solutions faster than the simplex-DEA approach.
3. EDEA is less sensitive to differences in problem characteristics, data quality, or tuning parameters than the simplex-DEA.
4. EDEA is easier to implement than the simplex-DEA.
5. EDEA has a wider range of applicability compared to the simplex-DEA.

For the experimental tests an input-oriented constant returns to scale model and data for the fiscal years 1991–1997 from the AHRP database were

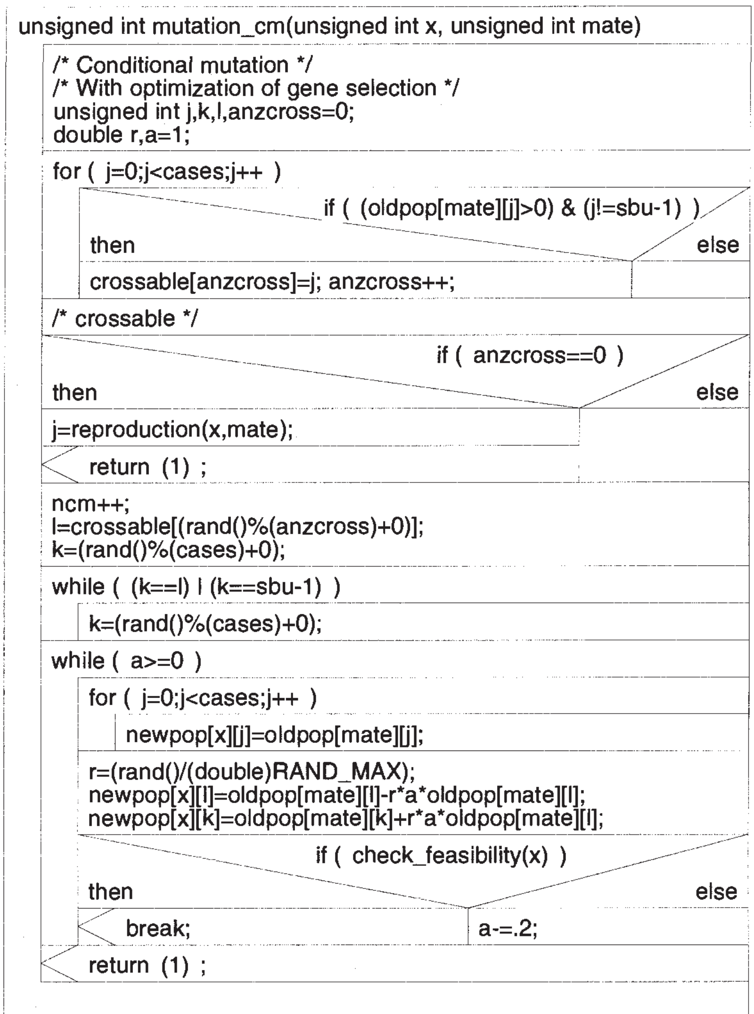


Fig. 8.18. Conditional mutation.

selected. The data for 61 Austrian hotels includes three output variables and four input variables; three of the input variables were defined as non-discretionary variables.

In order to obtain an optimal mixture of parameter settings for EDEA, several tests have been performed. The final settings, which were held constant during the EDEA runs, are illustrated in Table 8.5.

8.4.1 Does EDEA find similar solutions compared to LP-DEA?

For the EDEA/LP-DEA comparison, multiple runs were made under different values of parameters and the best result was chosen (the parameter settings in Table 8.5 refer to this best solution). In EDEA, the population size was 40 and the maximum number of generations was set to 100,000 throughout all trials.

There are basically two possibilities for comparing EDEA and LP-DEA solutions: (i) using the efficiency scores or (ii) using the recommended comparison partners.

Comparing efficiency scores

The comparative findings concerning the efficiency scores are summarized in Appendix Table A.11 and displayed in Fig. 8.19. Obviously there is a high correlation between the EDEA and LP-DEA findings. Some small deviations demand explanations.

Generally speaking, the EDEA efficiency scores for almost every company are slightly higher than the LP-DEA efficiency scores. Moreover, 29 companies (8.9%), which have been identified as inefficient by LP-DEA, are classified as best-practice companies in EDEA. This is either caused by EDEA not finding the global optimum (minimum) or by finding the global optimum in one

Table 8.5. Parameter specifications in EDEA test runs.

Parameter	Value	Description
pop_size	40	Population size
pm_um	0.02	Probability of uniform mutation
pm_bm	0.01	Probability of boundary mutation
pm_nm	0.02	Probability of non-uniform mutation
pm_cm	0.08	Probability of conditional mutation
pm_sc	0.10	Probability of simple crossover
pm_sa	0.10	Probability of single arithmetical crossover
pm_wa	0.10	Probability of whole arithmetical crossover
a	0.25	Coefficient used by the whole arithmetical crossover
b	2	Coefficient used for the simulated annealing function in the non-uniform mutation
fmultiple	4	Coefficient used by the scaling procedure

generation but then dropping it again in order to search for an even better solution (which does not exist).

The Pearson correlation coefficient between EDEA and LP-DEA results are 0.970 (sign. < 0.01), and 0.960 (sign. < 0.01) when excluding the efficient companies from the evaluation. Note that the LP-DEA results for efficient companies are achieved by running an additional (slightly modified) DEA model, whereas in EDEA the scores for efficient companies are calculated in one optimization run.

Comparing benchmarking partners

The comparison of the efficiency scores calculated by LP-DEA and EDEA have resulted in small differences between the two methodologies. But what about the recommended benchmarking partners? For the selection of comparison partners it is less important that the overall efficiency scores are exactly the same. In this situation, it is more important that the comparison of the recommended benchmarking partners leads to similar managerial recommendations. Furthermore, a manager who wants to benchmark his/her company may be more interested to learn about the superimposed (virtual) company, which offers insights into potential improvements for the company under evaluation.

By comparing the lambda values calculated by both methodologies analysis of EDEA and LP-DEA benchmarking partners can best be performed. The Pearson correlation coefficient between EDEA and LP-DEA lambda values

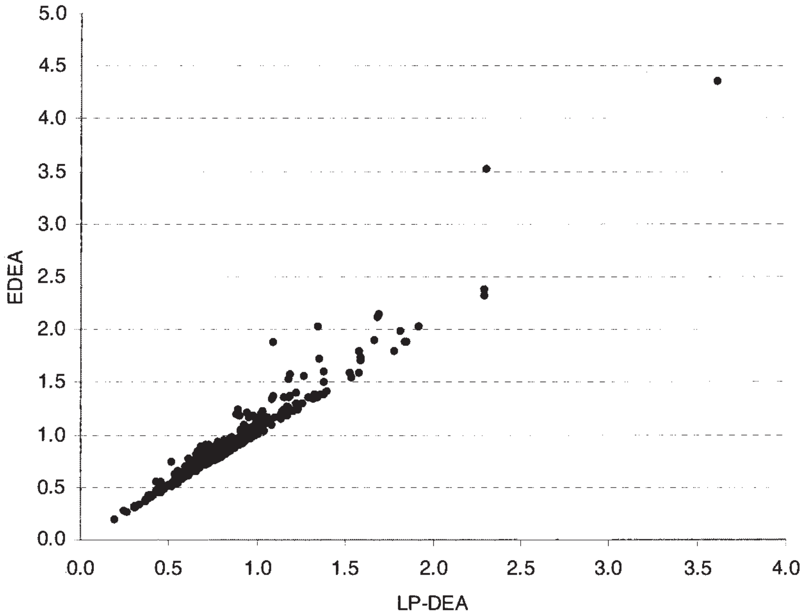


Fig. 8.19. Comparing EDEA and LP-DEA efficiency scores.

is 0.699 (sign. < 0.01), indicating a relatively poor relationship when considering the high correspondence of the efficiency scores (see p. 184). In fact, the recommendations of 'best-practising' partners evaluated by the two methodologies can lead to very different sets of peers. An example for this quite different selection of peers is illustrated in Fig. 8.20 and Table 8.6.

Figure 8.20 compares the EDEA and LP-DEA peers for hotel no. 396 in 1995. The overall inefficiency scores for hotel no. 396 calculated by EDEA and LP-DEA are 0.40 and 0.39, respectively. The LP-DEA analysis indicates that hotel no. 396 has the potential to improve its gross profit by €526,000 when it reduces it expenditures to €331,000 (a decrease of 61%), whereas the EDEA analysis indicates that it can improve its gross profit by €514,000 when reducing its expenditures to €343,000 (a decrease of 60%). Although both programs show almost identical scores and target values, the recommended comparison partners and their lambda values are very different.

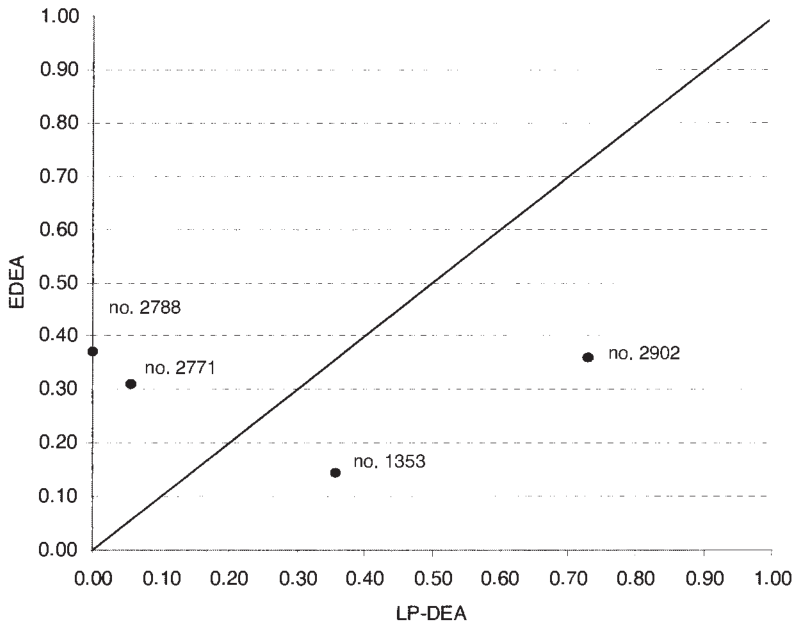


Fig. 8.20. Comparing EDEA and LP-DEA peers for hotel no. 396 (1995).

Table 8.6. Comparing EDEA and LP-DEA peers for hotel no. 396 (1995).

	Hotel					
λ	no. 1353	no. 2044	no. 2099	no. 2771	no. 2788	no. 2902
LP-DEA[a]	0.358			0.056		0.731
EDEA[b]	0.142	0.001	0.001	0.310	0.371	0.360

[a]e_{396} = 0.39; [b]e_{396} = 0.40 after 100,000 iterations.

Moreover, hotel no. 2788, which was identified as a significant benchmarking partner for hotel no. 396 by EDEA, was not considered in the LP-DEA solution. The small difference of the inefficiency scores between EDEA and LP-DEA certainly does not justify completely ignoring hotel no. 2788 as a potential benchmarking partner for hotel no. 396.

The phenomenon of similar target values generated by very different sets of comparison partners can best be explained by the different characteristics of the two optimization techniques. In DEA, input targets are calculated by the peers' input values weighted by the corresponding lambda values. In traditional linear programming these parameters are derived deterministically and lead to a single optimal solution. In EDEA the lambda values are generated by an approximation procedure and continuously improve during run-time. Figure 8.21 shows an example for the development of peer members for hotel

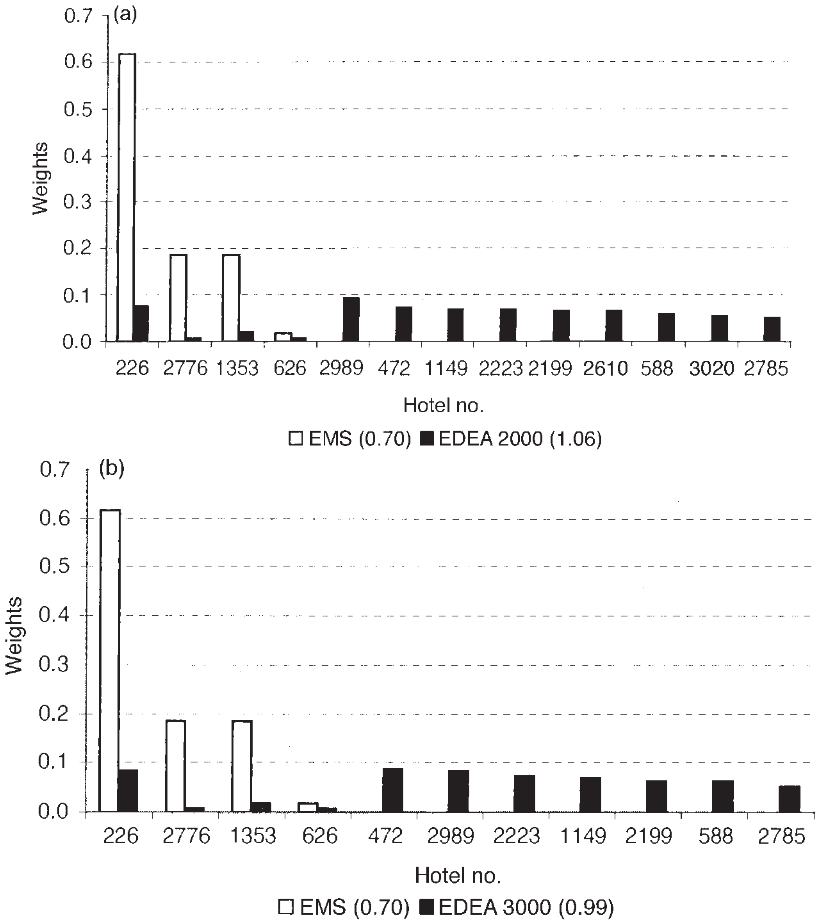
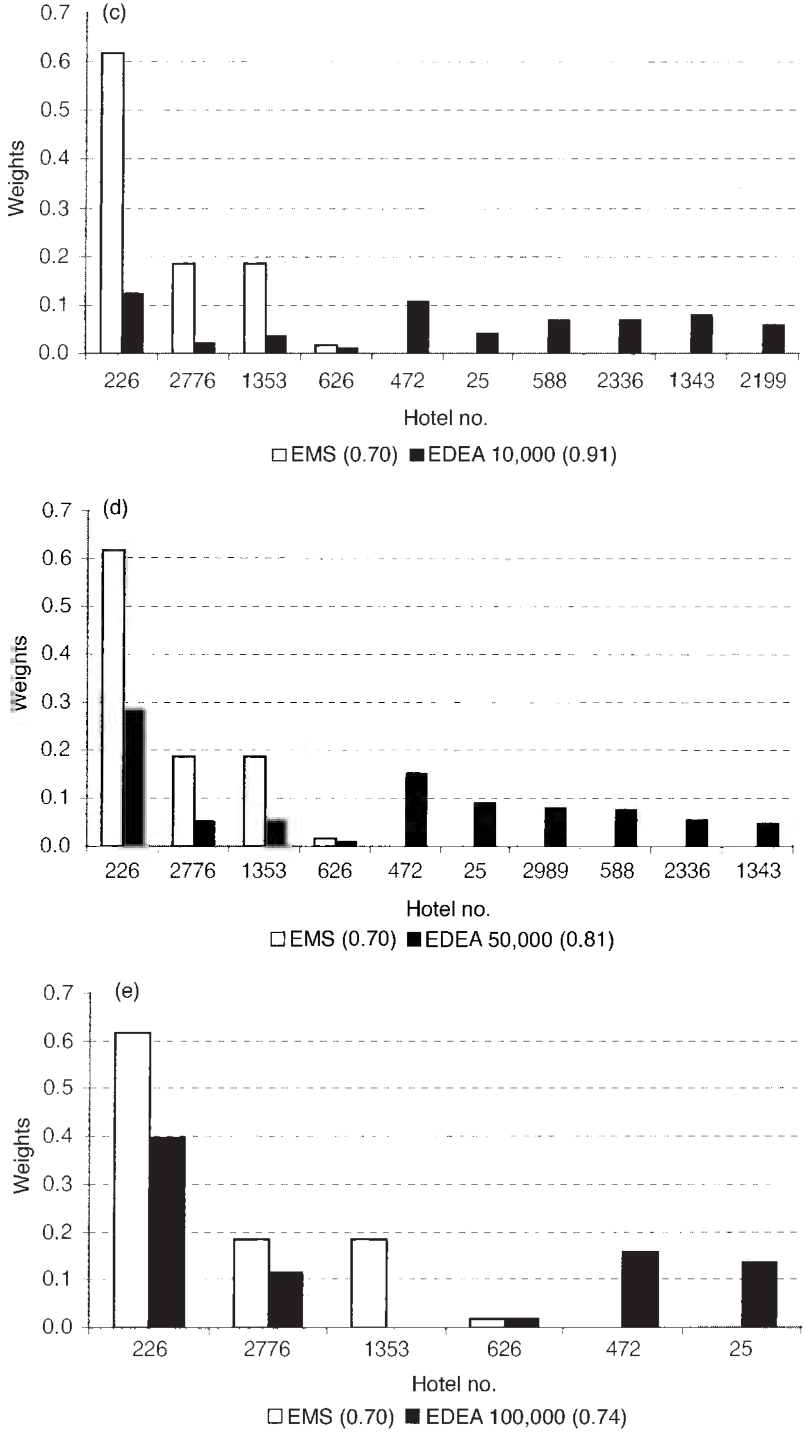


Fig. 8.21. Development of peer members for hotel no. 803 ($\lambda > 0.05$). (a)–(e) Five different periods during an optimization run.


(c)
Weights
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
226 2776 1353 626 472 25 588 2336 1343 2199
Hotel no.
EMS (0.70) EDEA 10,000 (0.91)
(d)
Weights
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
226 2776 1353 626 472 25 2989 588 2336 1343
Hotel no.
EMS (0.70) EDEA 50,000 (0.81)
(e)
Weights
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
226 2776 1353 626 472 25
Hotel no.
EMS (0.70) EDEA 100,000 (0.74)

no. 803 from the data set (fiscal year 1997). The five diagrams show peer members with $\lambda > 0.05$ for five different periods during an EDEA optimization run. (For a better comparison the optimal set of peer members, calculated by LP-DEA (EMS), are superimposed in the diagram.)

In the beginning of the optimization run, after 2000 iterations, when the EDEA efficiency score is still far away from the general optimum (0.7 vs. 1.1), the peers are very different compared to the optimal set of peers from the LP solution. Ten members have achieved a lambda value above 0.05 indicating their importance for the company under evaluation. Two hotels (no. 2776 and no. 1353), which are important peers in the optimal solution, have only marginal weights in the EDEA solution. After 3000 generations the efficiency score in EDEA falls below 1, thus classifying hotel no. 803 as an inefficient company. Companies no. 2610 and no. 3020, which were important members after 2000 generations, have disappeared from the set of significant comparison partners. In generation 10,000, when the efficiency score of hotel no. 803 dropped to 0.91, the number of peer members with $\lambda > 0.05$ declined to seven. Compared to the previous evaluation, some new members became important (no. 25, no. 2336, no. 1343) whereas others were eliminated from the 'consideration set' of comparison partners (no. 2989, no. 2223, no. 1149 and no. 2785). After 50,000 generations the picture of peer members is still very different compared to the optimal set derived by the LP-DEA algorithm. Note that the inefficiency score is 0.81 and is not far away from the optimal value. Finally, after 100,000 generations the peer structure significantly adjusts to the optimal set of comparison partners. It is remarkable that the elimination of a considerable number of potential comparison partners takes place during the last generations before program termination. Continuing the EDEA process would sooner or later come to an almost identical set of comparison partners compared to the LP-DEA solution.

What is remarkable from the evolutionary approach proposed here is that the process of the investigation of benchmarking partners can be substantially improved by also considering the alternative solutions generated during EDEA runtime. In addition to knowing the optimal set of peer members, alternative partners can offer important insights in business processes, not only when some of the optimal comparison partners turn out to be unsuited. In real managerial and economic problems, there are certainly situations where some of the 'near-optimal solutions' are more preferable than those solutions optimizing individual objectives. Furthermore, the number and diversity of the alternative benchmarking sets associated with a specific company often reflects the stability of a company's efficiency scores relative to the changes in efficiency criteria. Generally speaking, the more alternative sets of benchmarking partners that appear during the optimization run, the more sensitive the company's efficiency scores to the changes in criteria, the less reliable the company's efficiency scores. In summary, EDEA is softening the deterministic and immoveable type of solution generated with traditional LP-DEA algorithms by proposing alternative potential benchmarking partners.

8.4.2 Does EDEA produce solutions faster than LP-DEA?

When comparing EDEA and LP-DEA, the fact that the EDEA system allows the user to provide the desired number of generations, which influences the precision of the result, must be stressed. Nevertheless, comparing the runtime, LP-DEA clearly outperforms EDEA. Whereas EDEA needed approximately 10 min for one complete set of efficiency evaluations, DEAP required only 30 seconds. This finding is in line with other comparative studies where, considered in time, GA was outperformed by simplex-based algorithms (e.g. Dorsey and Mayer, 1995: 64).

One additional note in this context should be made here. Figure 8.22 shows the development of the efficiency score for one of the hotels in the data set. The illustration, which is very typical for all EDEA runs, clearly shows that major improvements in the optimization process are achieved during the first 25,000 generations. After a relatively short time EDEA can determine an approximate efficiency score, which might be sufficient for most practical purposes. The time consuming part in EDEA is the fine-tuning of the scores as illustrated in Fig. 8.22. There is no doubt, however, that more advanced fine-tuning GA operators that are currently being developed (e.g. Michalewicz and Fogel, 2000: 277) will significantly improve the performance of EDEA in the future.

8.4.3 Is EDEA less sensitive to problem characteristics and parameters?

As a result of the probabilistic nature of the genetic algorithm, it may be necessary to make multiple independent runs of the algorithm in order to obtain a satisfactory result for a given problem.

Also the relatively large number of parameters raises questions about the suitability of the genetic approach. In fact, the performance of an EDEA depends largely on the appropriate settings of initial parameter values. Moreover, during the course of a run, the optimal value for each operator probability may vary. Only the non-uniform mutation utilizes a dynamically variable value, depending on the age of the population. However, this imposes a fixed schedule which is problematic when different problems of varying complexity should be solved with EDEA.

There are several proposals for dynamic parameter adaptation in the GA literature that could be considered for future improvements of EDEA. The first suggestion was made by Booker (1987) who utilized a dynamically variable crossover rate depending on the spread of the fitnesses. When the population converges, the crossover rate is reduced to give more opportunity for mutation in order to find new variations. Davis (1991) describes another adaptive technique that is based directly on the success of an operator at producing good offspring. Credit is given to each operator when it produces a chromosome better than any other in the population. A weighting figure is allocated to each operator, based on its performance over the past 50 iterations. The operator is selected probabilistically, according to the current set of operator weightings. Therefore, during the course of a run, operator probabilities vary in an adaptive, problem-dependent way. The advantage of this technique is that it allows new operators to be compared directly with existing ones. However, a potential drawback of this technique is that it may reward operators which simply locate local optima rather than help find the global optima.

8.4.4 Is EDEA easier to implement than LP-DEAs?

There are several distinct differences between the working principles of GAs and other traditional optimization methods. The main differences compared to traditional optimization methods are related to the ability of GAs to work with a population of solutions instead of a single solution. Also, it does not require any other auxiliary information except the objective function values.

Another difference with most of the traditional methods is that GAs are suitable for parallel machines which reduces the overall computational time substantially. Considering the fast-growing computer industry and the significant achievements in software technology in the last decade, it is relatively undisputed that GA will become even more important in the future.

There are considerable opportunities for applying the EDEA approach in real-time environments. In contrast to the simplex approach, the adaptive characteristics of the genetic algorithm allow continuous learning from new observations. For instance, because in DEA only a few best-practising companies appear for the majority of comparative runs, it is obvious that the starting

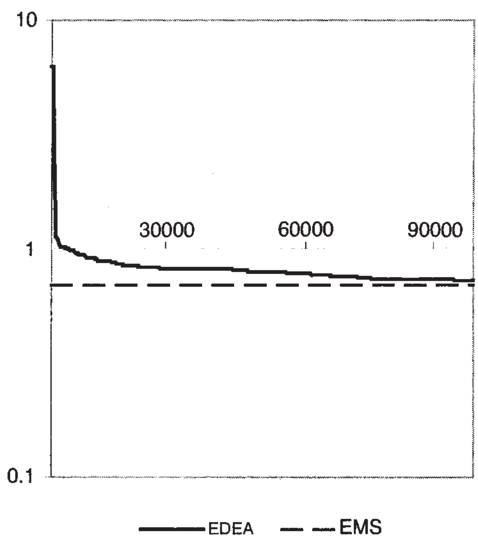


Fig. 8.22. Development of an efficiency score for hotel no. 803 ($\lambda_{EMS} = 0.696$; $\lambda_{EDEA} = 0.726$; gen = 100,000).

values of the weight vector may contain information from the last optimization run. In fact, an efficient real-time decision support system, like the AHRP benchmarking system implemented within the TourMIS[5] system, may use randomly generated starting values only once, at the very first time. Additionally, this extension to the basic algorithm will considerably save computation time and therefore reduce some of the performance gap between EDEA and simplex-based DEA.

8.4.5 Does EDEA have a higher applicability than LP-DEAs?

In EDEA, during the execution of the program, the user can request a current solution which always obeys the constraints, whereas LP-DEA is featured with an 'all or nothing' characteristic. This permits EDEA to offer alternative sets of comparison partners during run-time, which could be an important source of information for the user.

As discussed before, for the optimal selection of comparison partners it might not only be important to locate the global optima that are found with traditional DEA applications but also other near optima solutions. EDEA can give answers to the question 'What are other good sets of "best-practising" partners?' by prompting comparative partners with high weight loadings once a certain threshold has been reached (e.g. when $e < 1$). This capability can be used to guarantee that a potential benchmarking partner is not overlooked only because it is slightly dominated by another company in the data set.

Another major advantage of EDEA relates to the treatment of efficient companies in the analysis. In traditional LP-DEA, a value of 1 is assigned to efficient companies and nothing else can be said except that they are efficient. Although the extension of Andersen and Petersen (1993) seems to overcome this limitation, their approach is deficient for two reasons. First, the procedure involves a reformatting of the original DEA model with a restarting of as many DEA runs as efficient companies have been detected in the original run. Second, as discussed by several authors, and proven in the present study, the Andersen and Petersen adaptation leads in many cases to infeasible solutions, where ordinary linear programming algorithms break down. Beside the additional computational burden caused by redefining and rerunning a DEA model, the procedure is also not very practical and this may be one of the reasons why only a few authors have applied this approach in practice.

The EDEA procedure directly derives scores for efficient companies that can be interpreted in the way suggested by Andersen and Petersen (1993). For example, recall the simple input-oriented DEA example from Coelli (1996: 16) pictorialized in Fig. 8.1. The results for this DEA problem generated by EDEA and DEAP are summarized in Table 8.7.

---

5 tourmis.wu-wien.ac.at

The important difference illustrated here is that EDEA generates scores even for efficient units, which are the companies P_2 and P_5 in this example. For both companies the efficiency score reaches 2, which can be checked by the graphical representation of the problem in Fig. 8.23. For instance, the elimination of company P_2 in the spanning of the reference set implies that company P_2 is compared to that inefficient point in the input possibility set spanned by the remaining set of observations represented by the thick solid line in Fig. 8.23. The reference point thus becomes P_2'. In analogy to the inefficiency index, the efficiency index is calculated by $\overline{0P_2'}$, the line from 0 to P_2', divided by $\overline{0P_2}$, the line from 0 to P_2, which is exactly 2, the value found by EDEA. Similar to P_2 the analyses are performed for the other efficient company P_5. In this case the 'new' frontier is represented by the dashed line in Fig. 8.23.

EDEA can generate scores for efficient and inefficient companies in one round of computation, whereas DEA models based on the simplex algorithm are constrained to getting the results for efficient companies by performing multiple runs. This is another advantage compared to the traditional methodology which will raise the applicability of DEA models in general, and may support the development of DEA-based decision support systems in particular.

Table 8.7. Ranking of efficient companies in EDEA.

	DEAP						EDEA					
DMU	e	λ_1	λ_2	λ_3	λ_4	λ_5	e	λ_1	λ_2	λ_3	λ_4	λ_5
1	0.500		0.500				0.500		0.499			
2	1.000		1.000				2.000			0.339		
3	0.833		1.000			0.500	0.833		0.999			0.499
4	0.714		0.214			0.286	0.714		0.214			0.286
5	1.000					1.000	2.000		0.395	0.145	0.773	

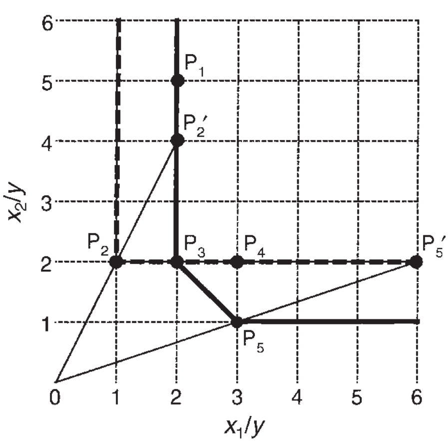


Fig. 8.23. Efficiency scores for companies P_2 and P_5.

Chapter 9

Perspectives of Benchmarking Decision Support Systems in Tourism and Hospitality Management

9.1 Turning Experience into Breakthrough Results

For many years, firms and organizations have borrowed ideas and compared themselves to others in hopes of discovering new ways to compete. In the 1960s, this was called cheating; in the 1970s, it was named reverse engineering. Today people call it benchmarking. However, when thoroughly designed and applied, benchmarking is much more than just finding out how an organization compares. As a methodological, well-prepared effort to investigate and discover best practices it can turn measures into breakthrough results.

In tourism, benchmarking is considered to have enormous potential in the quality improvement of services, particularly in the hotel and restaurant sector. That is, internal and external benchmarking technologies may help hotels and restaurants identify problems or patterns in the quality of services provided to customers. This can lead to changes in core accommodation and F&B processes (e.g. reservation, manufacturing processes) that improve customer service strategies. Anecdotal evidence in the tourism and hospitality literature documents the relationship between benchmarking and quality improvement; however, little work has been done to isolate this relationship empirically or analytically.

So far this book has described the state-of-the-art of benchmarking methodologies, in particular the selection of benchmarking partners, and how to extract managerial implications from analytical findings. It has discussed the various types of data-driven analytical techniques that represent the current supply of decision aids, which are necessary for building more intelligent benchmarking decision support systems in the future.

Although highly demanded by the industry and intensively discussed in the tourism and hospitality literature, the number of IT supported benchmarking systems is still very small. The fields of tourism and hospitality management and benchmarking support systems are not static, however. There are four principal issues which must be addressed when someone wants to take a closer look at the perspectives of benchmarking decision support systems in tourism and hospitality management.

1. The level of standardization in the tourism industry.
2. The development of process-related financial and non-financial performance measures.
3. The availability of analytical capabilities which are necessary to support the benchmarking process.
4. The development of information technology in general.

One can be sure that by 2010 the state-of-the-art of benchmarking decision support systems will be very different from what it is today. This implies that the requirements for the systems will also change. At the same time, progress on the information technology side are constantly offering new possibilities for the development of decision support systems. This final chapter reflects on these issues and perspectives and discusses their implications for building intelligent on-line benchmarking decision support systems in the future.

9.1.1 Standardization

Maybe with the exception of generic benchmarking, the objectives of all other benchmarking examinations request comparable data sources. Comparative studies in tourism are difficult to perform because of problems with accurately comparing products, measures and survey methodologies.

Market research in tourism is usually based on accommodation statistics, results from sample surveys of guests, accommodation providers or other experts or estimates achieved by analysing other statistical sources. Similar data sources can be typically found in the hospitality sector where additional information can be derived from financial statements and other internal cost and revenue control systems.

The two main difficulties tourism managers are faced with when comparing tourism data within and between tourism organizations are availability and comparability. Comparative data on financial statements of small and medium-sized companies and on tourism demand across Europe do not exist because of the different methodologies and/or sample bases used from one region (or one company) to another. In particular, survey data, frequently used for measuring non-financial performance factors, vary significantly in terms of instrument and methodology between different companies. The same is true with qualitative measures used for the assessment of tourism in destination management. In city tourism, for example, surveys used to estimate the importance of tourism sometimes ignore day-trippers or excursionists who usually generate a significant share of travel to cities and towns. The result is that urban tourism demand in these places is considered to be grossly underestimated which makes meaningful comparisons very difficult.

The methodological recommendations for tourism surveys provided by national and international tourism organizations, and the accounting rules which exist for hospitality businesses inevitably represent the statistical aims of governments. Regardless of the importance of standards put forward by international organizations (e.g. the Uniform Systems of Accounts developed by the American Hotel and Motel Association), they always depend on the willingness of national governments or authorities to implement these standards.

Benchmarking with other tourism firms or destinations is a slow and cautious process that requires mutual trust, something that can be enhanced with the formalization of the relationship at an early stage. Even when a benchmarking partnership is established, comparability of key performance measures are only valid if the organizations concerned can be sure that like is being compared with like.

9.1.2 Measurement

You cannot manage what you cannot measure. The majority of benchmarking projects fail as they do not define performance operationally. A company's quest for continuous improvement requires the use of quantitative data for problem solving, decision making, action planning and change. Developing a strong database and gathering from it as much relevant information as possible to plan organizational and operational improvements must be of high priority. The management approach in the tourism and hospitality sector has moved from manager-centred to customer-centred and the emphasis in a modern business context is delivering quality rather than producing quantity. For this reason hotels are more and more surveying employees and customers to determine satisfaction levels.

Modern business productivity is based on people productivity. It is therefore essential that performance measures are linked to reward and recognition systems. Non-financial performance measures, if implemented properly, are a very good morale booster and motivator. People like to know how well they are doing and where they should focus their attention for improving further. Measures can be used to show a standard and to establish comparisons.

For example, at its Courtyard and Residence Inn brands, Marriott has implemented a pay plan that also focuses on teams. The plan allows employees to earn up to 10% of their base pay in a bonus per quarter if the hotel reaches its profitability goals and the employee's team meets goals regarding:

- hotel guest satisfaction, measured by guest exit surveys;
- hotel productivity, measured in revenue per occupied room; and
- targeted team/hotel goals.

The targeted goals may be set by each individual hotel or by the guest care (front desk or restaurant) or the room care (housekeeping) teams. The teams set the goals based on what they collectively believe requires the most improvement, such as increasing check-out speed, improving breakfast service or improving room cleanliness.

The team is also responsible for evaluating whether it achieved its goals. Their input helps ensure that employees who contribute to the team are not overlooked by hotel management. Likewise, it helps to ensure that employees that do not complete their fair share of the work are not unjustly rewarded.

Nowadays hospitality businesses have almost no choice but to use customers as quality-control agents. However, the costs and benefits of guest involvement in the assessment of services need to be reviewed regularly to find the break even point where incremental costs are outweighed by incremental benefits.

9.1.3 Methodology

Benchmarking will benefit tremendously from developments in fields such as econometrics, statistics and operations research. The more data exist on the business side (or likewise in tourism destination marketing), the more sophisticated techniques are needed. The lack in technical details provided by the benchmarking literature and the misunderstandings frequently found in practice, have highlighted the need for a unified approach to capturing and analysing benchmarking data. A key step in benchmarking processes is the comparative analysis of key metrics to establish what constitutes 'best practice', the standard against which all others are compared. Armed with the most primitive data analysis tools, today's benchmarking analysts have no structured means to evaluate the data, characterize and measure performance gaps and project future performance levels. Case examples presented in this book have shown that the benchmarking process can be significantly enhanced by new, unifying data-analysis methodologies. Basic analytical models have been presented that formalize certain intuitive assumptions about competitive environment, firm productivity and profitability. The models introduced here attempt to support the selection of benchmarking partners and to develop 'best case scenarios' for benchmarking processes. From all the methodologies discussed throughout, Data Envelopment Analysis (DEA) has been identified as the most powerful technique which fulfils many of the requirements related to a good benchmarking endeavour.

Although theoretical production functions are inherently unknowable, empirical production functions, or efficient frontiers, can be constructed from observed data by DEA. DEA analyses each decision-making unit separately and measures its relative efficiency with respect to the entire set of units being evaluated. Further, this approach:

1. Emphasizes best practice, rather than distance-from-average practice, as with regression;
2. Does not require parametric assumptions about the underlying data relationships;
3. Handles various different assumptions about returns-to-scale;
4. Provides metrics of inefficiency for those units that are not exhibiting best practice;
5. Permits incorporation of user preferences into the analysis; and
6. Suggests routes of improvement for inefficient performers.

Compared to other methodologies DEA can cope more readily with multiple inputs and multiple outputs. However, the simplex algorithm traditionally applied in DEA models occurs as relatively inflexible and deterministic, which can hinder automatic adaptations and also prevents the detection of (non-optimal) alternative sets of comparison partners. These alternative sets of potential benchmarking partners can, nevertheless, be significantly better than the company under evaluation, and therefore contain important information for the benchmarking process. EDEA, the evolutionary counterfeits to the LP-based DEA models, can be employed to detect these alternative sets.

Extensive tests of LP-based DEA models and EDEA have shown that the proposals of benchmarking partners made by EDEA and LP-DEA runs indicated a relatively poor correspondence. However, for the purpose of benchmarking, it is less important that the proposed sets of benchmarking partners are identical. From a practitioner's point of view, it is more important that the comparison of recommended benchmarking partners leads to similar managerial recommendations. Although, the individual benchmarking partners were partly very different, both methodologies led to almost identical efficiency scores and target values.

What must be considered as an advantage of the evolutionary approach over the LP-based DEAs is that this new form of investigating comparison partners can substantially improve the benchmarking process. In addition to knowing the optimal set of peer members, alternative partners will certainly offer additional insights in business processes. Another advantage of EDEA compared to LP-DEA is that the former can generate scores for efficient and inefficient companies in one throw, whereas LP-DEA models based on the simplex algorithm are constrained by performing multiple runs for both groups.

Several advantages result when techniques supporting the selection of benchmarking partners are incorporated in an interactive environment. For example, for measuring a company's performance and competitiveness the input and output variables must be carefully selected to make the analysis useful for the manager. Although DEA has fewer limitations in the choice of input and output variables than the mixture regression approach, introduced earlier, the efficiency measure obtained by DEA is very sensitive to the combination of inputs and outputs. Thus, the input–output variables in a model are usually being based on intuitive or pragmatic considerations. The advantage of an interactive system is that the user can go back and forth and learn from the output. The manager will soon realize that the results may vary significantly, sometimes even through minor changes in the selected variables. He/she will learn from the system's response and gain a better understanding of how to interpret benchmarking results generated by the program. New perspectives for a more intensive use of interactive DEA applications occur in a multi-user environment like the Internet offers to SME hotel managers. Extranet applications of on-line databases of financial and non-financial hotel data offer an adequate platform for competitive analysis, especially for SMEs which are less organized in the exchange of business data than international (multiple unit) hotel chains.

Finally, the advantage of a real-time benchmarking application is that additional insights can be gained by multiperiod analysis and extrapolations of business data time series. In reality, the behaviour underlying the production process is likely to be dynamic because capital-intensive companies like hotels may take more than one period to adjust their choice variables to desired levels. Furthermore, capital inputs have a multiperiod dimension since they generate outputs in future periods. A decision support system based on longitudinal benchmarking data could therefore convert a performance evaluation model from an *ex post* evaluation instrument into a prospective oriented instrument. Overall, the present study has revealed that frontier techniques seem to be more suitable to be integrated in an interactive benchmarking system compared to central tendency techniques. However, practical problems in this context arise when a new company's data are added to the database. The deterministic characteristics of the standard LP-DEA model make it difficult to implement this methodology in modern decision support systems. EDEA, the evolutionary DEA approach proposed in this book, certainly overcomes some of these restrictions and is therefore an interesting alternative for future developments in this field.

9.1.4 Information technology support

Without doubt, tourism is one of the most popular arenas on the Internet and the economical impact within the next few years promises to be tremendous. It is, therefore, not very surprising that the first initiatives building online benchmarking decision support systems have started in the tourism and hospitality sector. These online benchmarking software applications are designed to prompt managers of their efficiency weaknesses based on programming that takes into account the hotel's productivity, environmental conditions and the main competitor's performance.

Starting in the more advanced accommodation sector, future benchmarking decision support systems will be better integrated in the daily operations and information flow of a company. One of the most important challenges in this context is to develop interfaces that can exchange information with the hotels' property management systems. This would enable benchmarking decision support software to monitor in real time all core business processes and automatically alert the manager when projected efficiency levels fall below a certain threshold. The key to such software's ability to aid management in optimizing their processes is the interface with computerized systems of other companies and the willingness of at least a minimum number of accommodation suppliers to share information on a continuous basis.

In addition, benchmarking software that interfaces with a property management system eliminates the need to ever rekey data from one software package to another, thus saving management both time and money. Managers can then choose from an arsenal of reports that they need to make well-informed business decisions. The selected reports could be formatted by employee, by cost centre or by business process. By comparing performance levels the reports will be goal-directed and support the achievement of the enterprise's overall priorities. It would encourage all users of the system to learn from other participants and supply insights to drive creative thinking. Hence, an online benchmarking decision support system supports the transformation of a company into a constantly knowledgeable, learning organization.

The objective of this book was to develop a powerful and practical framework for benchmarking decision support systems, and to validate their usefulness in a challenging service sector application. It suggests a quantitative foundation for the benchmarking process, grounded in mathematical and economic theory, that provides the capability to identify benchmarking partners and to analyse performance gaps.

In closing, benchmarking has been hotly debated within and outside the academic community with many explanations having been put forth. Throughout this book the author has deliberately simplified the complex benchmarking process with the goal of isolating the effect of selecting the optimal benchmarking partner on the performance evaluation of a firm. The analytical models presented in this book formalize certain intuitive assumptions about input and output relationships and environmental conditions usually experienced by accommodation and F&B providers. The author believes that future applications of these models will help to build intuition about the mechanisms driving these relationships.

Appendix A

Mixture Regression Models

Table A.1. Model Y_1 (F&B revenue).

Independent		Coefficient estimates	STD.ERR	*T*-value
#*x*	5	1		
X_1		0.140637	0.358924	0.391830
X_2		0.071471	0.113978	0.627056
X_3		0.084951	0.163549	0.519426
X_4		0.184613	0.151668	1.217217
X_5		0.454339	0.300119	1.513864
X_6		–0.100905	0.919370	–0.109755
X_7		1.141518	1.341685	0.850809
X_8		–0.102388	0.366164	–0.279625
X_9		0.054895	2.178933	0.025193
X_{10}		–0.874369	0.836804	–1.044890
X_{11}		–0.233099	0.942806	–0.247240
X_{12}		0.002382	0.055639	0.042812
Intercept		–0.172510	0.576720	–0.299123
Class size for segment		1 = 0.326911		
Variance for segment		1 = 0.021500		
#*x*	5	2		
X_1		0.467449	1.384456	0.337641
X_2		–0.976695	0.736035	–1.326969
X_3		2.113539	1.366304	1.546903
X_4		0.113504	0.207630	0.546664
X_5		0.165233	0.341655	0.483625
X_6		0.411483	1.875482	0.219401

Independent		Coefficient estimates	STD.ERR	T-value
X_7		0.404822	2.390658	0.169335
X_8		0.369251	0.473779	0.779373
X_9		1.238895	5.176412	0.239335
X_{10}		0.351353	0.699421	0.502349
X_{11}		–0.653444	1.209026	–0.540471
X_{12}		–0.085280	0.119511	–0.713581
Intercept		–5.523252	4.394873	–1.256749
Class size for segment		2 = 0.082107		
Variance for segment		2 = 0.183904		
#x	5	3		
X_1		–0.076980	0.202694	–0.379786
X_2		–0.016309	0.103132	–0.158133
X_3		0.067697	0.192518	0.351638
X_4		0.034653	0.158197	0.219050
X_5		0.222644	0.332367	0.669874
X_6		–0.019258	0.519846	–0.037045
X_7		–0.030652	0.693451	–0.044202
X_8		–0.001387	0.160631	–0.008637
X_9		0.261147	3.283089	0.079543
X_{10}		–0.212526	0.672363	–0.316089
X_{11}		0.057257	0.347159	0.164930
X_{12}		0.007931	0.084767	0.093556
Intercept		1.188832	0.758019	1.568341
Class size for segment		3 = 0.212946		
Variance for segment		3 = 0.002891		
#x	5	4		
X_1		–0.028276	0.436118	–0.064835
X_2		–0.046101	0.205490	–0.224345
X_3		–0.271989	0.324607	–0.837903
X_4		0.110042	0.149483	0.736149
X_5		0.148258	0.233145	0.635902
X_6		–0.044320	0.593632	–0.074658
X_7		0.405718	0.892248	0.454715
X_8		0.043804	0.305959	0.143168
X_9		0.007980	1.074458	0.007427
X_{10}		–0.102429	0.465343	–0.220115
X_{11}		–0.138504	0.429841	–0.322221
X_{12}		0.002567	0.062052	0.041373
Intercept		2.033814	1.209929	1.680937
Class size for segment		4 = 0.194018		
Variance for segment		4 = 0.004751		

Continued

Table A.1. *Continued.*

Independent		Coefficient estimates	STD.ERR	*T*-value
#*x*	5	5		
X_1		–0.078685	0.825036	–0.095372
X_2		–0.057993	0.232391	–0.249551
X_3		–0.074119	0.300468	–0.246680
X_4		–0.201171	0.164395	–1.223708
X_5		0.117599	0.372425	0.315764
X_6		2.237556	1.405006	1.592559
X_7		–0.006342	0.680839	–0.009315
X_8		0.062203	0.332430	0.187115
X_9		2.595941	3.547307	0.731806
X_{10}		0.603812	0.996270	0.606073
X_{11}		0.329619	1.247138	0.264301
X_{12}		–0.014582	0.073173	–0.199282
Intercept		0.963837	1.068625	0.901941
Class size for segment		5 = 0.184018		
Variance for segment		5 = 0.046177		

Table A.2. Model Y_2 (room revenue).

Independent		Coefficient estimates	STD.ERR	*T*-value
#*x*	5	1		
X_1		0.123311	0.949120	0.129921
X_2		–0.155331	0.325319	–0.477473
X_3		0.420587	0.742766	0.566244
X_4		0.102883	0.333396	0.308591
X_5		0.501128	0.709596	0.706216
X_6		–1.226714	2.438734	–0.503013
X_7		–1.147016	1.699285	–0.674999
X_8		–0.397171	0.838508	–0.473664
X_9		0.748313	2.418621	0.309396
X_{10}		–0.010570	1.114714	–0.009482
X_{11}		0.056166	1.721269	0.032630
X_{12}		0.103828	0.103541	1.002773
Intercept		–0.237957	2.671929	–0.089058
Class size for segment		1 = 0.131321		
Variance for segment		1 = 0.049666		
#*x*	5	2		
X_1		0.315502	0.200126	1.576514
X_2		0.057679	0.099321	0.580739
X_3		–0.202638	0.183220	–1.105984
X_4		0.049487	0.114047	0.433918
X_5		0.031948	0.264333	0.120862

Independent		Coefficient estimates	STD.ERR	*T*-value
X_6		–0.141346	0.359370	–0.393317
X_7		0.227910	0.655630	0.347620
X_8		0.036037	0.175385	0.205472
X_9		0.534149	2.169361	0.246224
X_{10}		–0.021592	0.433955	–0.049757
X_{11}		–0.000504	0.366851	–0.001373
X_{12}		–0.006532	0.062034	–0.105298
Intercept		1.119300	0.535020	2.092073
Class size for segment		2 = 0.262537		
Variance for segment		2 = 0.013266		
#*x*	5	3		
X_1		1.318430	0.858332	1.536037
X_2		0.090538	0.274877	0.329377
X_3		0.081498	0.328704	0.247938
X_4		0.072747	0.191436	0.380005
X_5		–0.205112	0.242915	–0.844378
X_6		0.461097	1.160661	0.397271
X_7		0.057412	1.617031	0.035504
X_8		0.231851	0.409386	0.566339
X_9		–0.052459	3.155743	–0.016623
X_{10}		–0.334415	0.968064	–0.345447
X_{11}		0.005312	1.041571	0.005100
X_{12}		0.014570	0.067484	0.215903
Intercept		–0.601504	1.038867	–0.579000
Class size for segment		3 = 0.191400		
Variance for segment		3 = 0.018129		
#*x*	5	4		
X_1		2.852609	0.861127	3.312646
X_2		–0.039307	0.163464	–0.240463
X_3		0.489431	0.318080	1.538703
X_4		–0.028461	0.255331	–0.111469
X_5		–0.086603	0.480802	–0.180122
X_6		–0.516292	1.603615	–0.321955
X_7		0.359136	0.638966	0.562058
X_8		–0.286443	0.855364	–0.334879
X_9		3.922726	2.989889	1.311997
X_{10}		1.192287	2.552600	0.467087
X_{11}		–0.252309	1.645888	–0.153296
X_{12}		0.075927	0.076392	0.993914
Intercept		–3.146243	1.090996	–2.883827
Class size for segment		4 = 0.185479		
Variance for segment		4 = 0.084740		

Continued

Table A.2. *Continued.*

Independent	Coefficient estimates	STD.ERR	*T*-value
#*x* 5	5		
X_1	0.073404	0.353613	0.207583
X_2	–0.031976	0.146866	–0.217725
X_3	–0.221389	0.187664	–1.179709
X_4	0.065491	0.126042	0.519595
X_5	0.064158	0.248079	0.258618
X_6	0.030847	0.618518	0.049872
X_7	0.417864	0.973842	0.429088
X_8	0.018763	0.156194	0.120129
X_9	1.008383	2.431756	0.414673
X_{10}	0.173456	0.436711	0.397188
X_{11}	–0.133982	0.308930	–0.433697
X_{12}	0.004786	0.066203	0.072289
Intercept	1.589596	0.646353	2.459330
Class size for segment	5 = 0.229262		
Variance for segment	5 = 0.007648		

Table A.3. Model Y_3 (average annual bed occupancy).

Independent	Coefficient estimates	STD.ERR	*T*-value
#*x* 3	1		
X_1	–0.187399	0.427354	–0.438511
X_2	–0.054269	0.119696	–0.453390
X_3	–0.130603	0.216091	–0.604386
X_4	0.076324	0.144470	0.528304
X_5	–0.108725	0.276793	–0.392802
X_6	–0.224045	1.051763	–0.213019
X_7	0.105532	1.140575	0.092526
X_8	0.056745	0.291743	0.194503
X_9	–0.366034	2.512784	–0.145669
X_{10}	–0.211123	0.497618	–0.424267
X_{11}	0.176837	0.754736	0.234303
X_{12}	–0.026622	0.061959	–0.429668
Intercept	1.445958	0.815402	1.773306
Class size for segment	1 = 0.235692		
Variance for segment	1 = 0.003987		
#*x* 3	2		
X_1	0.004080	0.275666	0.014799
X_2	–0.043064	0.134596	–0.319950
X_3	–0.163205	0.213674	–0.763802
X_4	0.019614	0.087875	0.223206
X_5	0.002883	0.166152	0.017350

Independent		Coefficient estimates	STD.ERR	T-value
X_6		–0.003318	0.431801	–0.007684
X_7		–0.157313	0.650377	–0.241879
X_8		–0.002295	0.161510	–0.014207
X_9		0.184885	0.795565	0.232394
X_{10}		–0.017668	0.345195	–0.051182
X_{11}		0.028679	0.292284	0.098119
X_{12}		0.011020	0.044441	0.247966
Intercept		0.889132	0.719925	1.235034
Class size for segment		2 = 0.366465		
Variance for segment		2 = 0.007318		
#x	3	3		
X_1		–0.110450	0.155104	–0.712100
X_2		–0.038800	0.077446	–0.501002
X_3		–0.010057	0.132756	–0.075753
X_4		–0.004527	0.100894	–0.044864
X_5		0.092585	0.155080	0.597016
X_6		0.029147	0.593109	0.049143
X_7		–0.209868	0.377965	–0.555259
X_8		0.013959	0.145818	0.095731
X_9		0.264827	1.782960	0.148532
X_{10}		–0.002026	0.445435	–0.004548
X_{11}		–0.011303	0.296863	–0.038074
X_{12}		0.005712	0.046869	0.121863
Intercept		0.630520	0.491137	1.283795
Class size for segment		3 = 0.397843		
Variance for segment		3 = 0.007491		

Table A.4. Efficiency scores of the mixture regression models.

	e_t	e_1			e_2			e_3			e_4			e_5			e_6			e_7		
j	no.	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3
1	14	–0.02	–0.22	–0.18	0.07	–0.09	0.01	0.11	0.16	0.17	0.03	–0.04	0.10	0.11	–0.24	0.02	0.14	–0.39	–0.05	0.07	–0.26	0.02
2	25	0.10	–0.10	–0.14	0.18	0.06	–0.02	0.16	0.09	–0.01	0.10	0.04	–0.09	0.08	–0.10	–0.09	0.14	–0.01	0.09	0.08	–0.24	–0.18
3	38	–0.08	0.05	–0.12	–0.06	0.11	0.22	–0.19	0.09	0.11	–0.13	0.13	0.23	–0.21	0.00	–0.03	–0.24	–0.05	–0.10	–0.33	–0.07	–0.10
4	42	0.06	0.08	0.06	0.03	0.15	0.07	0.02	0.08	–0.03	0.07	–0.22	–0.04	0.04	–0.24	–0.12	–0.15	0.05	0.01	–0.02	0.10	0.02
5	81	0.13	–0.01	0.18	0.04	–0.06	0.01	0.07	–0.12	–0.07	0.06	–0.05	–0.09	0.01	–0.10	–0.05	0.04	–0.01	0.08	0.12	0.06	–0.20
6	226	0.07	0.11	0.04	0.03	0.01	–0.04	–0.05	0.00	–0.11	–0.12	0.00	–0.05	0.00	–0.07	–0.11	0.07	0.06	0.03	0.09	–0.02	0.02
7	252	–0.23	0.01	0.05	–0.07	0.10	0.00	–0.05	0.17	0.05	–0.17	0.00	–0.04	–0.03	–0.21	0.03	0.16	–0.38	0.08	–0.13	–0.29	0.08
8	283	–0.02	0.10	–0.06	0.04	–0.07	–0.04	0.01	–0.06	0.09	0.06	–0.02	0.12	–0.01	–0.01	0.04	–0.07	–0.10	–0.02	0.03	0.20	–0.12
9	378	0.05	–0.05	0.14	0.01	–0.02	0.14	–0.06	0.02	0.06	–0.10	0.05	0.07	0.00	0.06	–0.05	–0.09	0.05	–0.08	–0.05	–0.02	–0.12
10	396	–0.04	–0.34	–0.05	–0.07	–0.50	–0.21	0.00	–0.13	0.02	0.05	–0.13	0.14	–0.06	0.04	0.39	–0.02	–0.15	–0.09	–0.11	–0.06	–0.31
11	419	–0.06	–0.14	–0.30	0.09	–0.41	–0.30	–0.13	–0.81	–0.10	–0.28	–0.03	0.04	–0.11	0.22	–0.21	–0.30	0.00	–0.10	–0.19	–0.14	0.17
12	472	0.14	–0.12	0.06	–0.07	–0.14	0.03	0.06	–0.09	0.00	0.02	0.00	–0.08	–0.07	–0.12	0.04	0.02	0.04	–0.03	–0.05	0.16	0.18
13	588	–0.03	0.18	–0.11	0.63	0.01	–0.04	–2.7	0.07	–0.19	–2.0	0.34	–0.21	0.16	–0.26	–0.10	0.30	–0.30	–0.04	0.22	–0.15	–0.06
14	626	0.11	0.14	0.05	–0.70	–0.02	0.05	–0.10	–0.01	–0.07	0.45	–0.15	–0.16	0.04	–0.11	–0.02	0.18	–0.09	0.06	–0.18	0.04	–0.07
15	688	0.04	0.05	0.09	–0.07	0.03	0.02	–0.15	–0.07	–0.17	0.01	0.05	–0.04	0.14	0.06	–0.17	0.11	–0.04	0.13	0.10	–0.01	–0.23
16	732	0.06	–0.15	0.11	–0.02	–0.07	0.02	–0.16	0.08	0.07	–0.15	–0.03	0.09	–0.32	0.02	0.11	–0.32	0.02	0.07	–0.16	–0.20	–0.17
17	803	–0.01	0.02	0.05	0.00	0.13	–0.39	0.00	0.04	–0.10	–0.08	–0.10	–0.18	–0.09	–0.28	–0.11	–0.08	–0.06	0.04	–0.11	0.01	0.25
18	834	–0.02	0.04	–0.20	–0.01	–0.08	–0.15	0.07	0.01	–0.21	–0.02	0.06	–0.05	0.05	0.08	–0.11	0.04	0.11	–0.10	0.01	–0.12	–0.13
19	921	0.40	–0.67	0.20	–0.11	0.05	0.11	–0.14	0.25	0.09	–0.12	–0.07	0.09	0.07	0.01	0.04	–0.18	0.10	0.08	–0.04	0.03	–0.10
20	1061	–0.49	0.08	–0.24	0.12	–0.05	–0.04	0.04	0.11	0.05	–0.17	0.25	–0.09	–0.10	–0.01	–0.02	–0.50	0.05	0.10	0.08	–0.03	–0.14

21	1128	0.55	–0.11	0.10	0.05	0.31	0.20	–0.50	0.17	–0.15	–0.28	–0.03	–0.34	0.01	–0.05	–0.56	–0.21	–0.30	–0.51	–0.41	–0.13	0.36
22	1134	0.21	0.19	0.06	–0.11	0.02	–0.15	–0.01	–0.03	0.04	–0.13	–0.14	–0.03	–0.40	0.10	0.03	–0.03	0.09	–0.08	0.23	0.09	–0.49
23	1149	–0.03	–1.0	–1.2	0.01	–0.51	–1.5	0.05	–0.42	–1.3	0.03	–0.03	–0.07	–0.04	0.15	0.04	–0.11	0.12	0.10	–0.17	0.09	0.28
24	1206	0.23	–0.09	0.20	0.24	–0.04	0.09	0.00	–0.09	0.05	0.13	0.14	0.15	0.07	–0.14	–0.08	0.14	0.23	–0.03	0.06	0.03	0.05
25	1246	0.02	–0.14	0.06	0.02	0.01	0.10	0.05	0.07	0.04	0.05	0.03	0.14	0.08	0.10	0.10	0.00	–0.05	–0.01	0.04	–0.07	0.12
26	1343	0.13	0.13	0.00	–0.25	0.45	–0.06	0.25	–0.48	–0.07	0.25	0.02	0.32	0.06	0.04	–0.08	0.28	–0.65	–0.03	0.23	–0.04	0.15
27	1353	–0.15	0.06	–0.02	–0.22	0.16	0.13	0.03	0.21	0.14	–0.07	0.15	0.12	0.23	0.09	–0.06	–0.22	–0.15	–0.22	0.23	0.17	0.01
28	1617	–0.08	–0.10	0.01	–0.02	–0.09	–0.09	–0.01	0.01	0.07	–0.06	–0.09	0.03	0.01	–0.04	0.03	0.09	0.14	0.07	0.03	0.09	0.17
29	1674	0.04	–0.06	0.02	0.01	–0.09	–0.05	0.04	–0.05	0.05	0.15	–0.07	0.11	0.11	–0.04	0.05	–0.04	–0.06	0.02	–0.02	–0.05	0.02
30	2005	0.12	0.04	–0.14	0.12	–0.01	–0.04	0.01	0.15	0.00	–0.25	–0.02	0.04	–0.15	–0.20	–0.06	0.00	–0.34	0.03	–0.55	0.08	–0.04
31	2042	0.08	–0.03	0.11	0.01	–0.03	0.03	–0.02	–0.03	0.04	–0.02	–0.02	0.09	–0.02	0.01	–0.12	0.01	–0.18	–0.15	–0.01	0.03	–0.08
32	2044	0.05	0.11	–0.23	0.24	0.21	0.07	0.20	–0.21	0.06	0.19	–0.26	0.06	0.07	–0.52	0.16	–0.14	–0.05	–0.03	0.19	–0.39	–0.04
33	2099	–0.38	–0.14	–0.07	0.00	0.20	–0.21	–0.15	0.23	0.04	0.06	–0.10	–0.17	0.16	–0.23	0.09	0.29	–0.08	0.18	0.14	0.15	–0.18
34	2199	–0.12	0.00	–0.01	–0.07	0.00	–0.06	0.17	–0.12	–0.03	–0.10	0.13	0.05	0.10	–0.05	0.00	0.16	–0.10	0.01	0.07	–0.04	0.08
35	2223	–0.06	–0.01	0.13	–0.10	–0.06	–0.05	–0.07	0.09	–0.01	0.21	0.13	–0.06	0.04	0.12	0.01	–0.03	0.12	0.05	0.10	0.15	0.08
36	2227	–0.03	–0.07	–0.02	0.00	–0.12	0.08	–0.09	–0.21	–0.23	0.10	0.03	0.01	0.11	0.09	–0.03	0.07	0.09	–0.05	0.00	0.07	0.02
37	2235	–0.02	0.14	–0.08	–0.05	0.13	0.00	0.00	0.18	0.12	0.04	0.07	0.15	0.09	–0.24	–0.02	0.01	0.08	0.02	–0.15	–0.08	–0.27
38	2280	0.03	–0.03	–0.20	0.03	–0.04	–0.14	–0.02	0.06	0.01	–0.01	0.04	0.00	0.02	0.01	0.06	0.05	–0.13	0.00	0.07	0.04	0.06
39	2287	–0.01	0.10	0.09	–0.15	0.14	0.16	–0.06	0.03	0.10	–0.15	–0.02	0.04	–0.01	–0.13	0.08	0.04	–0.09	–0.47	0.07	–0.13	–0.09
40	2336	0.07	–0.22	0.01	0.26	–0.23	0.12	–0.17	0.14	–0.04	–0.14	0.15	0.03	–0.19	0.04	0.01	–0.05	0.14	0.09	–0.17	0.14	0.05
41	2537	0.15	0.01	0.24	–0.10	0.03	0.17	–0.02	–0.06	0.11	–0.03	–0.04	0.09	–0.03	–0.08	0.04	–0.08	0.01	0.09	–0.07	0.03	0.07
42	2610	0.14	0.51	0.16	0.07	0.43	0.05	–0.04	–0.13	–0.90	0.04	0.32	–0.27	–0.11	–0.08	–0.64	–0.13	–0.17	–1.2	–0.25	–0.12	–0.74
43	2631	0.06	–0.04	0.17	0.07	–0.01	0.09	0.01	–0.11	0.03	–0.06	–0.16	–0.11	0.00	–0.12	0.04	–0.13	–0.26	–0.15	–0.07	–0.11	–0.07
44	2661	0.01	–0.01	–0.16	0.10	–0.10	–0.16	0.03	0.07	–0.03	–0.05	–0.10	–0.14	0.11	0.07	–0.07	0.28	–0.03	0.29	0.35	–0.15	0.31
45	2695	–0.04	0.35	–0.03	–0.16	0.26	–0.06	–0.20	0.20	–0.01	–0.07	0.12	–0.09	–0.05	0.05	0.01	–0.05	–0.05	–0.04	0.03	–0.20	–0.23

Continued

Table A.4. *Continued.*

	e_t	e_1			e_2			e_3			e_4			e_5			e_6			e_7		
j	no.	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3
46	2771	0.01	–0.05	–0.12	0.02	–0.17	0.04	–0.01	–0.15	–0.07	–0.04	0.17	–0.12	0.02	0.27	–0.03	–0.12	0.28	–0.52	–0.12	0.36	–0.48
47	2776	–0.11	–0.09	–0.01	–0.04	–0.04	–0.08	0.01	0.03	–0.16	0.01	–0.03	–0.08	0.08	–0.02	0.04	0.03	–0.09	–0.06	–0.01	–0.13	–0.06
48	2785	–0.04	–0.35	–0.69	–0.17	–0.12	–0.28	–0.13	0.03	–0.17	–0.03	0.06	–0.13	0.10	–0.01	–0.07	0.18	–0.02	–0.01	0.08	–0.15	–0.12
49	2788	–0.20	0.06	0.04	–0.03	–0.31	–0.27	–0.04	–0.19	0.27	–0.16	–0.25	–0.07	0.22	–0.08	0.18	0.42	0.03	–0.16	–0.64	–0.08	–0.33
50	2811	–0.06	0.02	0.17	0.01	0.05	0.02	0.02	0.09	0.09	–0.03	–0.19	–0.13	–0.03	0.09	–0.21	0.02	0.13	–0.04	–0.03	0.04	–0.06
51	2871	–0.01	0.11	0.09	–0.03	0.05	0.07	–0.06	–0.08	0.03	–0.01	0.11	0.12	0.06	0.05	–0.04	0.01	0.03	–0.11	0.02	–0.19	–0.12
52	2892	–0.05	0.07	0.23	–0.06	–0.05	0.09	0.02	–0.09	0.07	–0.06	–0.11	0.25	0.00	0.01	0.33	–0.03	0.05	0.35	–0.13	0.02	0.33
53	2902	–1.3	–1.1	–0.22	0.09	–0.25	0.18	–0.08	0.07	–0.07	0.18	0.29	0.30	–0.03	0.43	0.04	0.03	0.44	0.07	–0.14	0.21	–0.30
54	2914	–0.21	–0.64	–0.72	0.01	0.15	–0.05	–0.02	0.16	0.04	0.10	0.14	0.10	–0.05	0.12	0.09	–0.01	–0.01	0.07	0.02	–0.17	–0.10
55	2921	–0.02	–0.07	0.14	–0.17	–0.05	–0.12	–0.05	–0.05	–0.02	0.08	0.07	0.09	0.13	–0.09	0.09	0.18	0.12	0.09	0.09	0.02	0.07
56	2989	–0.30	0.45	–0.09	0.24	–0.42	–0.51	0.30	–0.60	–0.12	0.10	0.15	0.14	–0.23	–0.01	–0.47	–0.19	0.11	0.06	–0.36	0.21	0.15
57	2995	0.01	0.11	0.04	0.04	–0.12	0.09	0.08	–0.05	–0.23	–0.07	0.01	0.08	–0.06	0.14	0.23	0.14	–0.18	–0.31	0.12	0.02	–0.24
58	3004	0.22	0.06	–0.09	0.25	–0.21	–0.14	0.07	–0.04	–0.01	0.11	0.07	–0.43	–0.01	0.26	0.27	0.01	0.19	0.34	0.07	0.22	0.22
59	3020	–0.11	0.27	–0.12	–0.22	0.21	–0.24	0.10	0.22	0.02	0.22	0.32	–0.01	0.20	0.22	–0.01	0.12	0.07	–0.22	0.12	0.03	–0.03
60	3021	–0.08	–0.11	–0.18	0.07	–0.01	0.32	0.14	–0.07	–0.04	0.15	0.04	0.09	0.07	0.12	0.09	–0.01	–0.09	–0.09	–0.04	0.08	0.13
61	3044	0.00	0.12	0.14	–0.05	0.24	0.09	0.01	0.00	–0.25	0.00	–0.04	–0.03	–0.05	0.06	–0.10	0.04	0.13	0.03	0.01	0.01	0.04

Data Envelopment Analysis

Table A.5. Efficiency values of the basic VRS-DEA model.

		Efficiency values							Number of times a hotel is a peer							
n	no.	1991	1992	1993	1994	1995	1996	1997	91	92	93	94	95	96	97	91–97
1	14	0.847	0.974	1.000	1.000	0.997	1.000	1.000				1				1
2	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4	3	2	4	4	7	6	30
3	38	1.000	1.000	1.000	1.000	0.886	0.861	0.849			3	5				8
4	42	1.000	1.000	1.000	1.000	1.000	1.000	1.000	7	1	4	19	1	10	4	46
5	81	0.925	1.000	0.887	0.887	0.884	0.953	0.975		3						3
6	226	1.000	1.000	1.000	1.000	1.000	1.000	1.000	8	6	3	5	7	9	7	45
7	252	0.852	0.756	0.783	0.776	0.836	0.863	0.914								
8	283	1.000	1.000	1.000	1.000	1.000	0.948	0.915			6	3	1	0	0	10
9	378	1.000	1.000	1.000	1.000	1.000	1.000	1.000				3	6	2	7	18
10	396	0.628	0.624	0.640	0.577	0.597	0.477	0.467								
11	419	0.793	0.779	0.643	0.719	0.806	0.807	0.761								
12	472	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1			2		1		4
13	588	1.000	1.000	0.726	0.699	1.000	1.000	1.000	14	18				1	1	34
14	626	1.000	1.000	1.000	1.000	1.000	1.000	1.000	6	5	7	5	10	4	12	49
15	688	0.990	0.920	0.869	0.903	1.000	0.952	0.859								
16	732	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4	8	7	7	8	10	7	51
17	803	0.642	0.658	0.613	0.619	0.597	0.661	0.696								
18	834	0.922	0.830	0.840	0.823	0.724	0.842	0.752								
19	921	1.000	1.000	1.000	1.000	1.000	1.000	1.000			3	2	5	2	5	17
20	1061	0.948	0.921	1.000	1.000	0.930	1.000	1.000			4			6	11	21
21	1128	0.848	1.000	0.952	0.962	1.000	0.968	1.000		1						1
22	1134	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4	6	9	7	1	1	4	32
23	1149	0.820	0.801	0.787	0.755	0.735	0.808	0.807								
24	1206	0.866	0.940	0.853	0.994	1.000	1.000	1.000					2	9		11
25	1246	0.847	0.925	1.000	1.000	1.000	1.000	0.735			4	1				5
26	1343	1.000	0.910	1.000	1.000	0.800	1.000	1.000			2	2		1	1	6
27	1353	1.000	1.000	0.982	1.000	1.000	0.951	1.000					20		24	44
28	1617	1.000	0.915	0.915	0.886	0.856	0.889	0.848	1							1
29	1674	0.811	0.832	0.842	0.833	0.811	0.879	0.799								
30	2005	0.973	1.000	0.974	0.876	0.952	1.000	0.959		1						1
31	2042	0.824	0.869	0.833	0.869	0.881	0.734	0.736								
32	2044	0.766	0.982	0.830	0.857	0.862	0.814	0.868								
33	2099	0.864	0.943	0.979	0.966	1.000	1.000	0.977						1		1
34	2199	1.000	0.963	0.977	1.000	1.000	1.000	1.000				1	1		3	5
35	2223	1.000	0.890	1.000	0.938	0.954	0.964	0.924	8							8

Continued

Table A.5. *Continued.*

		Efficiency values							Number of times a hotel is a peer							
n	no.	1991	1992	1993	1994	1995	1996	1997	91	92	93	94	95	96	97	91–97
36	2227	0.770	0.795	0.742	0.863	0.843	0.806	0.751								
37	2235	1.000	1.000	1.000	1.000	0.933	1.000	0.997	19	21	25	6		9		80
38	2280	0.895	0.854	0.812	0.825	0.746	0.766	0.817								
39	2287	1.000	1.000	1.000	0.998	1.000	1.000	1.000					3		2	5
40	2336	1.000	1.000	1.000	1.000	1.000	1.000	1.000	10	2				1	5	18
41	2537	1.000	1.000	1.000	1.000	1.000	1.000	0.970	3	1			2	1		7
42	2610	0.782	0.791	0.689	0.775	0.666	0.723	0.774								
43	2631	1.000	1.000	1.000	1.000	1.000	1.000	1.000	18	11	9	13	15	10	14	90
44	2661	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2	5	5	3	3	5	7	30
45	2695	1.000	1.000	1.000	0.751	0.813	0.824	0.748	2		1					3
46	2771	1.000	1.000	1.000	1.000	1.000	1.000	1.000		7	4	21	14	15	16	77
47	2776	1.000	1.000	1.000	1.000	1.000	1.000	1.000	3	2	6	2	5	4	13	35
48	2785	0.990	0.996	1.000	1.000	1.000	1.000	1.000			10	10	1	4	4	29
49	2788	1.000	1.000	1.000	1.000	1.000	1.000	1.000	21	13	22	9	11	8	4	88
50	2811	0.763	0.837	0.847	0.724	0.835	0.987	0.833								
51	2871	0.713	0.711	0.648	0.795	1.000	1.000	0.753								
52	2892	0.934	0.807	0.792	0.790	0.755	0.785	0.755								
53	2902	1.000	1.000	1.000	1.000	1.000	1.000	1.000		14	1	10	6	6	1	38
54	2914	0.731	0.934	0.992	1.000	1.000	1.000	0.977				4	6	5		15
55	2921	0.867	0.741	0.855	0.942	0.836	0.905	0.827								
56	2989	0.758	0.786	0.847	0.880	0.844	0.859	0.886								
57	2995	0.840	0.815	0.804	0.811	0.799	0.745	0.773								
58	3004	0.875	0.864	0.893	0.747	0.954	1.000	1.000								
59	3020	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4	3	1	7	4		1	20
60	3021	0.846	1.000	0.665	0.717	0.877	0.770	0.777								
61	3044	0.740	0.791	0.685	0.708	0.699	0.796	0.697								

Table A.6. Window analysis results.

n	no.	91–1	91–2	91–3	92–1	92–2	92–3	93–1	93–2	93–3	94–1	94–2	94–3	95–1	95–2	95–3
1	14	0.84	0.93	0.96	0.94	1.00	0.99	0.98	0.97	1.00	0.98	0.99	1.00	0.99	1.00	0.97
2	25	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	0.95	1.00	0.98	1.00	1.00
3	38	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.88	1.00	0.87	0.83	0.86	0.83	0.83
4	42	1.00	1.00	0.98	1.00	0.98	1.00	1.00	1.00	0.90	1.00	0.90	1.00	0.92	1.00	1.00
5	81	0.88	0.91	0.86	0.91	0.86	0.81	0.86	0.81	0.80	0.88	0.86	0.89	0.88	0.91	0.94
6	226	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00
7	252	0.81	0.72	0.74	0.74	0.76	0.73	0.76	0.73	0.77	0.75	0.78	0.78	0.80	0.80	0.81
8	283	0.96	1.00	1.00	1.00	1.00	0.96	1.00	0.97	0.92	1.00	0.97	0.93	0.99	0.93	0.90
9	378	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.98	1.00
10	396	0.61	0.60	0.60	0.60	0.61	0.55	0.61	0.56	0.59	0.57	0.59	0.45	0.59	0.44	0.45
11	419	0.78	0.77	0.61	0.78	0.61	0.71	0.63	0.71	0.77	0.70	0.76	0.77	0.76	0.78	0.75
12	472	1.00	1.00	0.99	1.00	0.99	0.98	0.95	0.93	0.89	1.00	0.97	0.98	0.91	0.91	1.00
13	588	0.98	1.00	0.73	1.00	0.73	0.70	0.73	0.70	1.00	0.70	1.00	1.00	1.00	1.00	1.00
14	626	1.00	1.00	0.99	1.00	1.00	0.97	1.00	0.98	0.98	1.00	0.99	1.00	0.99	1.00	1.00
15	688	0.98	0.85	0.79	0.91	0.84	0.87	0.85	0.87	0.87	0.90	0.89	0.87	0.97	0.94	0.85
16	732	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
17	803	0.63	0.65	0.59	0.65	0.60	0.59	0.61	0.60	0.59	0.61	0.59	0.66	0.59	0.65	0.68
18	834	0.87	0.79	0.84	0.76	0.81	0.78	0.78	0.73	0.71	0.73	0.72	0.74	0.72	0.73	0.74
19	921	1.00	0.98	1.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	1061	0.93	0.85	1.00	0.83	1.00	0.93	1.00	0.93	0.86	1.00	0.90	0.98	0.89	1.00	1.00
21	1128	0.84	1.00	0.92	1.00	0.92	0.88	0.94	0.91	0.96	0.93	0.95	0.92	0.95	0.91	1.00
22	1134	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.84	1.00	0.98	0.91	1.00	0.95	1.00
23	1149	0.79	0.77	0.77	0.79	0.79	0.72	0.75	0.71	0.73	0.71	0.73	0.73	0.73	0.73	0.75
24	1206	0.83	0.89	0.85	0.88	0.85	0.96	0.81	0.88	1.00	0.88	1.00	1.00	1.00	1.00	0.96
25	1246	0.81	0.89	1.00	0.85	1.00	1.00	0.99	0.95	1.00	0.93	0.98	0.83	0.95	0.83	0.70
26	1343	0.93	0.90	1.00	0.75	0.96	0.98	0.88	0.97	0.76	1.00	0.75	0.86	0.75	0.84	1.00
27	1353	0.91	0.92	0.97	0.92	0.97	0.90	0.96	0.89	1.00	0.96	1.00	0.85	1.00	0.85	1.00
28	1617	0.89	0.88	0.90	0.86	0.87	0.85	0.85	0.83	0.82	0.83	0.82	0.78	0.82	0.78	0.83
29	1674	0.79	0.81	0.78	0.79	0.78	0.78	0.79	0.79	0.75	0.82	0.79	0.79	0.80	0.80	0.78
30	2005	0.94	1.00	0.95	1.00	0.92	0.86	0.92	0.86	0.86	0.88	0.87	0.93	0.88	0.95	0.96
31	2042	0.81	0.80	0.79	0.83	0.82	0.80	0.82	0.80	0.78	0.85	0.84	0.72	0.84	0.71	0.69
32	2044	0.74	0.94	0.82	0.91	0.82	0.82	0.81	0.81	0.86	0.81	0.86	0.78	0.83	0.78	0.79
33	2099	0.82	0.91	0.92	0.87	0.90	0.97	0.91	0.95	1.00	0.92	0.98	1.00	0.98	1.00	0.94
34	2199	0.93	0.93	0.97	0.94	0.97	0.94	0.97	0.93	0.94	1.00	0.98	0.95	1.00	0.98	0.98
35	2223	1.00	0.87	0.98	0.84	0.94	0.94	0.94	0.93	0.90	0.91	0.88	0.93	0.88	0.88	0.89

Continued

Table A.6. *Continued.*

n	no.	91–1	91–2	91–3	92–1	92–2	92–3	93–1	93–2	93–3	94–1	94–2	94–3	95–1	95–2	95–3
36	2227	0.74	0.76	0.71	0.76	0.72	0.81	0.71	0.80	0.82	0.79	0.81	0.78	0.81	0.79	0.73
37	2235	1.00	1.00	1.00	1.00	1.00	0.97	1.00	0.98	0.93	1.00	0.93	1.00	0.93	1.00	0.94
38	2280	0.87	0.83	0.80	0.79	0.76	0.79	0.74	0.77	0.74	0.77	0.74	0.72	0.73	0.71	0.74
39	2287	1.00	1.00	0.98	1.00	0.98	0.96	0.99	0.96	1.00	0.97	1.00	0.88	1.00	0.88	0.87
40	2336	1.00	1.00	0.96	1.00	0.94	0.93	0.97	0.97	0.97	0.96	0.97	1.00	0.96	1.00	1.00
41	2537	1.00	1.00	0.95	1.00	0.95	0.98	0.94	0.97	0.97	0.95	0.96	0.98	0.91	0.91	0.91
42	2610	0.75	0.77	0.67	0.75	0.67	0.73	0.67	0.73	0.66	0.74	0.66	0.70	0.67	0.71	0.74
43	2631	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	0.99
44	2661	1.00	0.99	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00
45	2695	1.00	0.91	0.97	0.91	0.93	0.74	0.91	0.74	0.73	0.74	0.74	0.75	0.75	0.76	0.74
46	2771	0.94	1.00	0.99	0.94	0.90	1.00	0.89	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
47	2776	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00
48	2785	0.92	0.95	1.00	0.94	0.98	1.00	0.99	1.00	1.00	0.99	0.98	1.00	0.98	1.00	0.98
49	2788	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.98	1.00
50	2811	0.75	0.77	0.81	0.79	0.83	0.69	0.82	0.69	0.74	0.72	0.77	0.93	0.79	0.95	0.83
51	2871	0.70	0.69	0.60	0.68	0.60	0.69	0.63	0.70	0.71	0.79	0.80	0.75	0.77	0.73	0.75
52	2892	0.89	0.81	0.74	0.81	0.75	0.73	0.76	0.73	0.74	0.74	0.75	0.74	0.75	0.73	0.71
53	2902	1.00	1.00	0.99	0.98	0.93	1.00	0.97	1.00	1.00	1.00	0.99	1.00	0.98	1.00	1.00
54	2914	0.72	0.93	0.95	0.89	0.91	1.00	0.92	1.00	0.99	1.00	1.00	0.99	1.00	0.99	0.96
55	2921	0.82	0.69	0.86	0.68	0.84	0.94	0.73	0.89	0.79	0.89	0.80	0.79	0.77	0.77	0.80
56	2989	0.75	0.78	0.83	0.78	0.83	0.84	0.83	0.85	0.83	0.86	0.84	0.84	0.84	0.84	0.84
57	2995	0.79	0.78	0.80	0.78	0.80	0.78	0.78	0.76	0.78	0.77	0.78	0.73	0.78	0.73	0.74
58	3004	0.83	0.84	0.87	0.84	0.88	0.70	0.88	0.70	0.75	0.71	0.77	0.91	0.88	1.00	0.86
59	3020	1.00	1.00	0.91	1.00	0.97	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
60	3021	0.79	1.00	0.66	1.00	0.65	0.71	0.59	0.68	0.82	0.68	0.82	0.69	0.80	0.66	0.78
61	3044	0.72	0.75	0.68	0.75	0.64	0.65	0.63	0.64	0.62	0.68	0.69	0.76	0.69	0.74	0.67

Table A.7. Window analysis peer weights.

n	*no.*	Mean	*N*	SD	Min	Max
1	14	0.277	10	0.228	0.012	0.609
2	25	0.190	51	0.189	0.001	0.783
3	38	0.221	68	0.246	0.001	0.943
4	42	0.172	122	0.123	0.001	0.686
6	226	0.167	121	0.164	0.001	0.829
8	283	0.161	38	0.208	0.001	0.794
9	378	0.233	52	0.263	0.007	0.966
12	472	0.401	10	0.377	0.035	0.920
13	588	0.148	85	0.179	0.001	0.936
14	626	0.185	149	0.221	0.001	0.968
16	732	0.189	177	0.156	0.001	0.686
19	921	0.187	44	0.246	0.006	0.984
20	1061	0.197	60	0.212	0.001	0.779
21	1128	0.112	2	0.073	0.060	0.163
22	1134	0.137	115	0.178	0.001	0.889
24	1206	0.188	90	0.159	0.001	0.733
25	1246	0.342	7	0.356	0.019	0.799
26	1343	0.157	13	0.121	0.017	0.413
27	1353	0.259	157	0.203	0.001	0.813
30	2005	0.113	10	0.115	0.007	0.332
33	2099	0.889	2	0.054	0.851	0.927
34	2199	0.256	8	0.329	0.011	0.842
35	2223	0.327	1		0.327	0.327
37	2235	0.289	244	0.223	0.001	0.918
39	2287	0.168	31	0.217	0.008	0.677
40	2336	0.150	16	0.175	0.005	0.609
41	2537	0.308	12	0.328	0.019	0.841
43	2631	0.230	254	0.200	0.002	0.977
44	2661	0.212	61	0.208	0.001	0.952
45	2695	0.257	8	0.288	0.001	0.816
46	2771	0.222	256	0.190	0.001	0.865
47	2776	0.119	91	0.177	0.001	0.942
48	2785	0.295	96	0.276	0.002	0.949
49	2788	0.275	285	0.206	0.003	0.802
53	2902	0.167	136	0.163	0.001	0.885
54	2914	0.199	71	0.215	0.001	0.860
59	3020	0.208	94	0.177	0.001	0.803

Table A.8. Malmquist efficiency, technical and productivity change.

	Efficiency change						Technical change						Productivity change					
no.	92	93	94	95	96	97	92	93	94	95	96	97	92	93	94	95	96	97
14	1.09	1.10	0.98	1.04	1.00	1.00	1.11	1.00	0.98	0.96	0.97	0.99	1.21	1.10	0.96	0.99	0.98	0.99
25	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.04	0.92	0.90	1.12	0.82	1.01	1.04	0.92	0.90	1.12	0.82
38	1.00	1.00	1.00	0.87	0.95	0.97	1.18	0.99	1.11	0.86	0.98	1.02	1.18	0.99	1.11	0.75	0.93	0.99
42	1.00	0.98	1.02	1.00	1.00	1.00	0.99	0.94	1.11	0.85	1.14	0.94	0.99	0.92	1.14	0.85	1.14	0.94
81	1.06	0.93	1.00	0.98	1.09	1.02	0.97	1.03	0.95	1.00	0.97	1.02	1.03	0.95	0.95	0.98	1.06	1.04
226	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.01	0.92	1.08	1.00	0.97	1.00	1.01	0.92	1.08	1.00
252	0.89	1.03	0.96	1.12	1.05	1.02	1.00	1.00	0.95	0.93	0.96	0.98	0.89	1.03	0.92	1.04	1.01	1.00
283	1.00	1.00	1.00	1.00	0.93	0.97	1.06	1.01	0.94	0.95	1.01	0.99	1.06	1.01	0.94	0.95	0.94	0.96
378	1.00	1.00	1.00	1.00	1.00	1.00	1.03	0.99	0.96	0.97	0.95	0.98	1.03	0.99	0.96	0.97	0.95	0.98
396	1.00	1.01	0.71	1.19	0.84	0.94	1.00	1.03	1.05	1.05	0.91	1.07	0.99	1.03	0.75	1.25	0.76	1.01
419	0.94	0.81	1.24	1.10	0.96	1.05	1.02	1.00	1.10	0.97	1.02	1.01	0.96	0.81	1.36	1.07	0.98	1.06
472	1.00	1.00	1.00	1.00	1.00	1.00	1.01	0.99	0.99	0.96	1.00	1.18	1.01	0.99	0.99	0.96	1.00	1.18
588	1.00	0.46	0.86	2.55	1.00	1.00	1.06	0.93	0.98	1.02	1.01	0.99	1.06	0.43	0.84	2.59	1.01	0.99
626	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	0.87	0.98	1.03	1.11	1.00	0.97	0.87	0.98	1.03	1.11
688	0.93	0.96	1.04	1.05	0.97	0.91	0.97	0.97	0.99	0.95	1.00	1.03	0.90	0.93	1.03	1.00	0.97	0.93
732	0.93	1.07	1.00	1.00	1.00	1.00	0.94	1.05	1.00	1.07	0.95	0.90	0.88	1.12	1.00	1.07	0.95	0.90
803	1.04	0.90	0.89	0.92	1.27	1.14	1.02	0.97	1.03	1.00	1.00	0.97	1.06	0.88	0.92	0.92	1.27	1.11
834	0.92	1.01	0.94	0.84	1.27	0.80	0.99	1.05	1.03	1.13	0.86	1.19	0.91	1.07	0.97	0.95	1.08	0.95
921	1.02	1.00	1.00	1.00	1.00	1.00	1.20	1.06	1.01	0.86	0.99	1.13	1.23	1.06	1.01	0.86	0.99	1.13
1061	0.94	1.15	0.99	0.92	1.10	1.00	1.00	1.02	0.97	0.99	0.96	1.99	0.94	1.17	0.96	0.92	1.06	1.99

1128	1.34	0.78	0.88	1.13	0.84	1.53	1.09	1.05	1.01	1.00	0.98	1.11	1.46	0.82	0.89	1.12	0.83	1.69
1134	1.00	1.00	1.00	1.00	1.00	1.00	0.97	0.94	0.61	0.43	0.93	1.03	0.97	0.94	0.61	0.43	0.93	1.03
1149	0.93	1.02	0.73	0.95	1.14	1.35	1.07	1.03	1.04	1.11	0.90	0.88	0.99	1.04	0.75	1.05	1.03	1.19
1206	1.10	0.91	1.12	1.05	1.00	1.00	1.00	1.05	1.04	1.65	0.67	0.96	1.09	0.96	1.16	1.73	0.67	0.95
1246	1.09	1.08	1.00	1.00	0.92	0.76	1.00	1.04	0.99	1.05	0.95	1.07	1.09	1.12	0.99	1.05	0.88	0.81
1343	0.89	1.12	0.97	0.83	1.17	1.01	1.08	1.01	1.04	0.98	0.96	1.13	0.95	1.13	1.01	0.81	1.12	1.14
1353	1.01	0.98	1.02	1.00	0.95	1.05	1.01	1.07	0.91	1.08	0.82	1.23	1.02	1.05	0.92	1.08	0.78	1.29
1617	0.93	1.00	0.92	1.01	1.03	0.97	1.06	1.02	1.03	0.99	0.92	1.10	0.98	1.03	0.95	1.00	0.95	1.06
1674	1.00	1.03	1.00	0.98	1.08	0.91	1.03	0.96	0.99	0.96	0.93	1.07	1.03	0.99	0.99	0.94	1.00	0.97
2005	1.03	0.97	0.89	1.03	1.07	0.96	1.08	0.92	1.07	0.97	0.99	1.07	1.11	0.90	0.95	1.00	1.06	1.03
2042	1.01	0.97	1.12	0.99	0.87	0.99	0.98	1.03	0.87	1.00	0.99	0.97	0.99	1.00	0.98	1.00	0.86	0.97
2044	1.33	0.83	0.94	1.09	0.78	1.15	1.01	1.04	1.04	1.13	0.89	1.01	1.34	0.86	0.98	1.24	0.70	1.16
2099	1.10	1.01	1.01	1.01	1.02	0.98	1.01	1.01	1.03	1.04	1.01	0.96	1.12	1.02	1.04	1.05	1.03	0.94
2199	0.97	1.02	1.03	1.00	1.00	1.00	1.04	1.04	0.92	0.94	0.93	1.02	1.01	1.06	0.95	0.94	0.93	1.02
2223	0.89	1.11	0.94	0.98	1.04	0.96	0.99	1.01	1.06	0.99	1.01	1.04	0.88	1.13	1.00	0.97	1.05	1.00
2227	1.02	0.88	1.19	1.00	0.95	0.88	1.01	1.04	1.04	1.02	1.02	1.06	1.03	0.92	1.23	1.01	0.97	0.93
2235	1.00	1.00	1.00	0.82	1.22	0.96	1.08	0.95	0.86	0.91	0.98	0.96	1.08	0.95	0.86	0.75	1.19	0.92
2280	0.94	0.96	1.02	0.90	1.03	1.03	1.01	1.01	1.02	1.06	0.92	1.05	0.96	0.97	1.04	0.96	0.94	1.08
2287	1.03	0.93	0.99	1.09	0.90	1.05	0.99	1.01	0.91	1.02	0.84	1.02	1.02	0.94	0.90	1.10	0.75	1.07
2336	1.00	1.00	0.96	1.04	1.00	1.00	1.03	0.88	1.02	0.94	1.12	1.00	1.03	0.88	0.99	0.98	1.12	1.00
2537	1.00	1.00	1.00	0.98	1.02	0.95	1.03	0.94	1.04	1.03	1.00	1.05	1.03	0.94	1.04	1.01	1.01	1.00
2610	0.94	0.76	1.18	0.69	1.01	1.12	1.04	1.01	1.08	1.09	0.93	1.05	0.97	0.76	1.27	0.75	0.94	1.18
2631	1.00	0.90	1.05	1.06	1.00	1.00	0.96	0.92	0.93	1.01	0.88	1.08	0.96	0.84	0.97	1.06	0.88	1.08
2661	1.00	1.00	1.00	1.00	1.00	1.00	0.93	0.98	0.99	1.06	1.03	1.05	0.93	0.98	0.99	1.06	1.03	1.05
2695	0.98	0.99	0.77	1.08	1.01	0.90	0.89	1.04	1.09	0.93	1.03	1.06	0.87	1.03	0.84	1.01	1.04	0.95

Continued

Table A.8. *Continued.*

	Efficiency change						Technical change						Productivity change					
no.	92	93	94	95	96	97	92	93	94	95	96	97	92	93	94	95	96	97
2771	1.00	1.00	1.00	1.00	1.00	1.00	1.08	0.97	1.22	0.97	1.02	1.05	1.08	0.97	1.22	0.97	1.02	1.05
2776	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	1.02	1.08	0.94	0.94	0.99	0.97	1.02	1.08	0.94	0.94
2785	1.00	1.01	1.00	1.00	1.00	1.00	1.02	1.06	1.04	1.00	1.04	0.95	1.02	1.07	1.04	1.00	1.04	0.95
2788	1.00	1.00	1.00	1.00	1.00	1.00	0.96	1.45	0.71	1.53	0.69	0.84	0.96	1.45	0.71	1.53	0.69	0.84
2811	1.04	1.03	0.91	1.08	1.17	0.90	0.99	1.01	0.96	0.96	0.98	1.03	1.03	1.04	0.87	1.04	1.15	0.92
2871	1.01	0.91	1.25	1.26	0.99	0.67	0.98	0.96	0.94	0.79	0.94	1.30	0.99	0.87	1.17	1.00	0.93	0.88
2892	0.83	1.00	0.96	0.99	1.05	0.93	1.04	0.93	1.04	1.06	0.93	1.05	0.87	0.93	1.00	1.05	0.98	0.98
2902	1.00	1.00	1.00	1.00	1.00	1.00	1.24	0.88	1.45	0.97	1.12	0.85	1.24	0.88	1.45	0.97	1.12	0.85
2914	1.30	1.05	1.02	1.00	1.00	0.95	1.01	0.98	1.11	0.98	0.98	0.99	1.31	1.03	1.14	0.98	0.98	0.94
2921	0.86	1.15	1.12	0.85	1.08	0.94	0.98	1.06	0.99	1.09	0.86	1.17	0.84	1.21	1.11	0.92	0.93	1.10
2989	1.33	1.17	1.24	0.83	1.04	1.10	1.03	1.01	0.99	1.02	1.01	0.95	1.36	1.18	1.23	0.85	1.04	1.05
2995	0.93	1.08	0.93	0.92	1.03	1.03	1.03	0.97	0.98	1.14	0.95	1.01	0.96	1.05	0.91	1.04	0.98	1.04
3004	0.94	0.95	0.86	1.30	1.07	0.95	1.01	1.08	0.97	0.97	0.99	0.98	0.95	1.03	0.83	1.26	1.06	0.93
3020	1.00	0.86	1.03	1.14	0.92	1.09	0.90	0.95	1.05	0.91	0.95	1.01	0.90	0.81	1.08	1.04	0.87	1.09
3021	1.08	0.73	1.08	1.06	1.01	0.95	1.05	1.04	0.99	1.07	0.89	1.15	1.14	0.76	1.06	1.14	0.90	1.09
3044	1.06	0.87	1.04	0.99	1.14	0.85	0.99	1.03	0.95	1.02	0.96	1.04	1.05	0.89	0.99	1.01	1.10	0.88

Table A.9. Efficiency values of the VRS-DEA with non-discretionary input variables.

		Efficiency values							Number of times a hotel is a peer							
n	no.	1991	1992	1993	1994	1995	1996	1997	91	92	93	94	95	96	97	91–97
1	14	0.795	0.952	1.000	1.000	0.992	1.000	1.000						4		4
2	25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	2	2		4	3	2	14
3	38	1.000	1.000	1.000	1.000	0.759	0.639	0.622			1					1
4	42	1.000	1.000	1.000	1.000	1.000	1.000	1.000				22		5	2	29
5	81	0.925	1.000	0.847	0.835	0.790	0.913	0.953		4						4
6	226	1.000	1.000	1.000	1.000	1.000	1.000	1.000	3					1	1	5
7	252	0.839	0.653	0.699	0.583	0.705	0.775	0.904								
8	283	1.000	1.000	1.000	1.000	1.000	0.882	0.857			1					1
9	378	1.000	1.000	1.000	1.000	1.000	1.000	1.000				2	3	1	2	8
10	396	0.587	0.547	0.572	0.338	0.402	0.313	0.271								
11	419	0.645	0.584	0.435	0.566	0.640	0.535	0.539								
12	472	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1							1
13	588	1.000	1.000	0.350	0.198	1.000	1.000	1.000	7	19						26
14	626	1.000	1.000	1.000	1.000	1.000	1.000	1.000	8	4	4	5	9	5	11	46
15	688	0.990	0.920	0.869	0.901	1.000	0.952	0.859								
16	732	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1	3	1	1	5	4	2	17
17	803	0.554	0.621	0.477	0.404	0.321	0.496	0.479								
18	834	0.922	0.825	0.824	0.746	0.630	0.770	0.541								
19	921	1.000	1.000	1.000	1.000	1.000	1.000	1.000			4	1	3	1	8	17
20	1061	0.860	0.921	1.000	1.000	0.928	1.000	1.000			6			11	8	25
21	1128	0.662	1.000	0.872	0.838	1.000	0.630	1.000								
22	1134	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4	6	8	4	1	1	1	25
23	1149	0.749	0.620	0.689	0.492	0.460	0.504	0.609								
24	1206	0.815	0.928	0.843	0.990	1.000	1.000	1.000						8		8
25	1246	0.844	0.918	1.000	1.000	1.000	1.000	0.546			1					1
26	1343	1.000	0.908	1.000	1.000	0.693	1.000	1.000			11	3		3	5	22
27	1353	1.000	1.000	0.964	1.000	1.000	0.925	1.000					25		27	52
28	1617	1.000	0.902	0.915	0.811	0.761	0.828	0.791	2							2
29	1674	0.783	0.832	0.737	0.763	0.784	0.850	0.728								
30	2005	0.965	1.000	0.959	0.873	0.937	1.000	0.959						1		1
31	2042	0.824	0.869	0.832	0.860	0.867	0.692	0.588								
32	2044	0.700	0.978	0.773	0.738	0.811	0.488	0.596								
33	2099	0.819	0.939	0.967	0.962	1.000	1.000	0.959						5		5
34	2199	1.000	0.945	0.951	1.000	1.000	1.000	1.000				1	4		2	7
35	2223	1.000	0.890	1.000	0.934	0.940	0.949	0.880	10		1					11

Continued

Table A.9. *Continued.*

		Efficiency values							Number of times a hotel is a peer							
n	no.	1991	1992	1993	1994	1995	1996	1997	91	92	93	94	95	96	97	91–97
36	2227	0.754	0.752	0.669	0.794	0.707	0.670	0.541								
37	2235	1.000	1.000	1.000	1.000	0.811	1.000	0.993	26	26	26	4		2		84
38	2280	0.885	0.854	0.771	0.766	0.631	0.693	0.636								
39	2287	1.000	1.000	1.000	0.996	1.000	1.000	1.000					1		1	2
40	2336	1.000	1.000	1.000	1.000	1.000	1.000	1.000	17	1				1	4	23
41	2537	1.000	1.000	1.000	1.000	1.000	1.000	0.950	2	1			1	1		5
42	2610	0.737	0.669	0.504	0.572	0.397	0.370	0.401								
43	2631	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2	3		5	5	1	3	19
44	2661	1.000	1.000	1.000	1.000	1.000	1.000	1.000	5	5	6	4	3	3	6	32
45	2695	1.000	1.000	1.000	0.686	0.696	0.735	0.616	2		3					5
46	2771	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2	15	3	24	18	17	18	97
47	2776	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2		3				3	8
48	2785	0.988	0.995	1.000	1.000	1.000	1.000	1.000			7	8	2	14	8	39
49	2788	1.000	1.000	1.000	1.000	1.000	1.000	1.000	20	4	19	4	6	4	5	62
50	2811	0.763	0.837	0.847	0.714	0.814	0.987	0.828								
51	2871	0.700	0.707	0.592	0.744	1.000	1.000	0.520								
52	2892	0.923	0.764	0.691	0.646	0.595	0.656	0.469								
53	2902	1.000	1.000	1.000	1.000	1.000	1.000	1.000		21	5	23	13	17	11	90
54	2914	0.700	0.906	0.988	1.000	1.000	1.000	0.929				2	5	3		10
55	2921	0.867	0.741	0.855	0.942	0.836	0.904	0.823								
56	2989	0.245	0.452	0.517	0.578	0.422	0.378	0.378								
57	2995	0.715	0.628	0.702	0.616	0.470	0.583	0.565								
58	3004	0.819	0.746	0.684	0.621	0.948	1.000	1.000								
59	3020	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2			5	1	1	3	12
60	3021	0.788	1.000	0.665	0.708	0.877	0.747	0.777								
61	3044	0.653	0.731	0.686	0.618	0.588	0.715	0.462								

Table A.10. Ranking for efficient companies in the non-discretionary VRS-DEA.

j	no.	1991	1992	1993	1994	1995	1996	1997
1	14	< 1	< 1	1.064	1.055	< 1	1.359	1.132
2	25							
3	38	1.191				< 1	< 1	< 1
4	42	1.267	1.222	1.099		1.170		
5	81	< 1	1.000	< 1	< 1	< 1	< 1	< 1
6	226							
7	252	< 1	< 1	< 1	< 1	< 1	< 1	< 1
8	283						< 1	< 1
9	378							
10	396	< 1	< 1	< 1	< 1	< 1	< 1	< 1
11	419	< 1	< 1	< 1	< 1	< 1	< 1	< 1
12	472							
13	588			< 1	< 1			
14	626							
15	688	< 1	< 1	< 1	< 1	1.015	< 1	< 1
16	732							
17	803	< 1	< 1	< 1	< 1	< 1	< 1	< 1
18	834	< 1	< 1	< 1	< 1	< 1	< 1	< 1
19	921							
20	1061	< 1	< 1	1.074	1.037	< 1	1.098	
21	1128	< 1		< 1	< 1	1.095	< 1	
22	1134							1.641
23	1149	< 1	< 1	< 1	< 1	< 1	< 1	< 1
24	1206	< 1	< 1	< 1	< 1			1.059
25	1246	< 1	< 1	1.283	1.155	1.034	1.003	< 1
26	1343	1.086	< 1	1.040	1.880	< 1	1.019	1.231
27	1353	1.011	1.009	< 1	1.000	1.681	< 1	1.897
28	1617	1.044	< 1	< 1	< 1	< 1	< 1	< 1
29	1674	< 1	< 1	< 1	< 1	< 1	< 1	< 1
30	2005	< 1		< 1	< 1	< 1	1.005	< 1
31	2042	< 1	< 1	< 1	< 1	< 1	< 1	< 1
32	2044	< 1	< 1	< 1	< 1	< 1	< 1	< 1
33	2099	< 1	< 1	< 1	< 1	1.120	1.095	< 1
34	2199		< 1	< 1				
35	2223	1.063	< 1	1.010	< 1	< 1	< 1	< 1
36	2227	< 1	< 1	< 1	< 1	< 1	< 1	< 1
37	2235					< 1		< 1
38	2280	< 1	< 1	< 1	< 1	< 1	< 1	< 1
39	2287	1.064	1.090	1.026	< 1		1.185	1.142
40	2336	1.388		1.970	1.757	1.344		

Continued

Table A.10. *Continued.*

j	no.	1991	1992	1993	1994	1995	1996	1997
41	2537	1.184	1.158	1.036	1.103	1.051	1.124	< 1
42	2610	< 1	< 1	< 1	< 1	< 1	< 1	< 1
43	2631							
44	2661	1.479	1.602	1.716	1.745	1.775	2.222	2.154
45	2695	1.461	1.104	1.081	< 1	< 1	< 1	< 1
46	2771	1.016	1.262	1.100	1.593	1.771	2.007	2.022
47	2776							
48	2785	< 1	< 1	1.476		1.275	1.359	1.251
49	2788							
50	2811	< 1	< 1	< 1	< 1	< 1	< 1	< 1
51	2871	< 1	< 1	< 1	< 1			< 1
52	2892	< 1	< 1	< 1	< 1	< 1	< 1	< 1
53	2902	1.250	1.902	1.190	3.696	2.369	2.293	2.017
54	2914	< 1	< 1	< 1	1.338			< 1
55	2921	< 1	< 1	< 1	< 1	< 1	< 1	< 1
56	2989	< 1	< 1	< 1	< 1	< 1	< 1	< 1
57	2995	< 1	< 1	< 1	< 1	< 1	< 1	< 1
58	3004	< 1	< 1	< 1	< 1	< 1	1.112	1.038
59	3020			1.050		1.861		
60	3021	< 1		< 1	< 1	< 1	< 1	< 1
61	3044	< 1	< 1	< 1	< 1	< 1	< 1	< 1

Note: Empty cells indicate infeasible LP solutions.

Table A.11. EDEA and LP-DEA results for non-discretionary CRS-DEA.

		1991		1992		1993		1994		1995		1996		1997	
j	no.	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA
1	14	0.79	0.83	0.86	0.95	0.97	1.01	0.91	1.05	0.99	1.13	1.13	1.19	1.13	1.16
2	25														
3	38	1.05	1.17			1.20				0.74	0.86	0.62	0.73	0.61	0.78
4	42	1.24		1.19	1.37	0.95	1.16	1.61		1.15	1.25				
5	81	0.86	0.91	0.93	1.00	0.80	0.87	0.82	0.91	0.77	0.91	0.90	1.00	0.94	1.02
6	226														
7	252	0.75	0.81	0.65	0.65	0.68	0.87	0.55	0.65	0.67	0.74	0.74	0.81	0.71	0.79
8	283											0.87	0.96	0.82	0.83
9	378														
10	396	0.57	0.59	0.53	0.54	0.54	0.55	0.33	0.34	0.39	0.40	0.30	0.31	0.26	0.26
11	419	0.64	0.67	0.58	0.63	0.43	0.46	0.54	0.55	0.61	0.67	0.52	0.54	0.47	0.48
12	472														
13	588					0.33	0.34	0.19	0.20						
14	626														
15	688	0.95	0.99	0.86	0.94	0.80	0.83	0.82	0.91	0.87	0.96	0.86	0.95	0.70	0.76
16	732	1.21	1.23	0.90	0.94	1.15	1.21	1.17	1.18	1.85	1.89	1.53	1.59	1.29	1.36
17	803	0.53	0.54	0.58	0.61	0.45	0.45	0.39	0.42	0.31	0.32	0.44	0.49	0.31	0.32
18	834	0.89	0.94	0.78	0.84	0.80	0.80	0.74	0.82	0.61	0.64	0.77	0.88	0.53	0.53
19	921	0.96	0.99												
20	1061	0.86		0.86	0.89	1.05	1.11	0.97	1.19	0.83	0.96	1.04	1.17		

Continued

Table A.11. *Continued.*

		1991		1992		1993		1994		1995		1996		1997	
j	no.	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA	LP	EDEA
21	1128	0.63	0.68	1.00	1.02	0.72	0.73	0.58	0.61	0.67	0.68	0.45	0.45		
22	1134													1.58	1.59
23	1149	0.73	0.78	0.61	0.61	0.66	0.77	0.38	0.43	0.41	0.42	0.38	0.39	0.59	0.60
24	1206	0.81	0.83	0.92	0.96	0.83	0.84	0.94	1.02			1.22	1.41	0.99	1.13
25	1246	0.83	0.85	0.91	0.94	1.14	1.22	1.07	1.11	1.02	1.10	0.84	0.94	0.54	0.54
26	1343	1.08	1.34	0.85	0.85	1.00	1.02	0.92	1.09	0.67	0.70	0.91	0.92	0.82	0.88
27	1353	0.97	1.08	1.00	1.05	0.95	1.00	0.99	1.09	1.68	2.12	0.91	1.00	1.69	2.15
28	1617	0.98	1.04	0.90	0.91	0.91	0.93	0.76	0.79	0.73	0.77	0.78	0.85	0.71	0.75
29	1674	0.74	0.78	0.79	0.79	0.72	0.77	0.72	0.91	0.65	0.75	0.82	0.88	0.58	0.65
30	2005	0.96	0.98			0.96	1.02	0.76	0.89	0.79	0.85	0.91	1.03	0.81	0.94
31	2042	0.68	0.76	0.70	0.83	0.66	0.82	0.79	0.86	0.76	0.84	0.65	0.72	0.55	0.57
32	2044	0.70	0.75	0.97	0.98	0.77	0.78	0.69	0.74	0.80	0.81	0.49	0.51	0.56	0.56
33	2099	0.82	0.84	0.93	0.95	0.90	1.18	0.96	0.96	0.95	1.21	1.03	1.23	0.95	1.07
34	2199	0.96	0.99	0.94	0.98	0.94	0.96	1.08	1.09						
35	2223	1.03	1.06	0.87	0.88	0.98	1.03	0.91	0.97	0.85	0.89	0.90	0.99	0.88	0.91
36	2227	0.75	0.77	0.75	0.81	0.67	0.69	0.77	0.84	0.71	0.87	0.66	0.73	0.53	0.63
37	2235									0.75	0.82	1.15		0.91	1.04
38	2280	0.87	0.89	0.82	0.87	0.75	0.77	0.76	0.79	0.60	0.61	0.62	0.67	0.55	0.58
39	2287	0.96	1.00	1.04	1.04	0.91	0.93	0.88	0.93	1.58	1.80	0.78	0.93	0.88	1.20
40	2336	1.38	1.60			1.17	1.53	0.70	0.71	1.32	1.39				

41	2537	1.18	1.26	1.15	1.36	1.02	1.20	1.09	1.38	0.96	1.00	1.09	1.88	0.89	0.93
42	2610	0.72	0.78	0.66	0.78	0.50	0.53	0.57	0.62	0.39	0.43	0.37	0.39	0.40	0.40
43	2631			1.05		0.84	0.85	0.92	0.96			1.02	1.14	1.17	1.23
44	2661	1.34	1.36	1.38	1.51	1.38	1.39	1.22	1.29	1.32	1.34	1.40	1.41	1.80	
45	2695	1.34	2.03	0.96	1.00	0.95	1.01	0.64	0.67	0.66	0.69	0.66	0.73	0.62	0.70
46	2771	1.02	1.15	1.26	1.30	1.10	1.16	1.58	1.74	1.67	1.91	1.92	2.03	1.81	1.99
47	2776														
48	2785	0.99	1.00	0.99	0.99	1.47				1.27	1.57	1.35	1.72	1.23	1.24
49	2788			2.03								1.59	1.71	1.54	1.55
50	2811	0.62	0.74	0.65	0.76	0.69	0.71	0.60	0.70	0.65	0.77	0.86	0.99	0.71	0.90
51	2871	0.64	0.65	0.65	0.69	0.58	0.61	0.71	0.78			0.96	1.07	0.52	0.56
52	2892	0.92	0.93	0.76	0.76	0.67	0.72	0.63	0.67	0.57	0.58	0.61	0.68	0.47	0.50
53	2902	1.18	1.37	1.84	1.86	1.19	1.57	3.61	4.36	2.30	3.53	2.29	2.38	1.78	1.80
54	2914	0.70	0.72	0.91	0.96	0.97	1.05	1.33						0.89	1.25
55	2921	0.84	0.86	0.70	0.75	0.82	0.82	0.93	1.00	0.68	0.86	0.79	0.87	0.61	0.74
56	2989	0.24	0.27	0.45	0.46	0.52	0.52	0.55	0.55	0.41	0.42	0.38	0.39	0.36	0.37
57	2995	0.71	0.75	0.62	0.76	0.66	0.71	0.60	0.67	0.43	0.56	0.57	0.60	0.56	0.56
58	3004	0.81	0.84	0.73	0.75	0.68	0.84	0.60	0.61	0.70	0.82	0.85	0.96	0.72	0.84
59	3020	2.30	2.32	1.36	1.40	0.78	0.85	0.76	0.80	1.17	1.27	0.82	0.82	1.16	1.20
60	3021	0.79	0.84	0.84	0.96	0.63	0.65	0.67	0.75	0.68	0.73	0.66	0.85	0.57	0.60
61	3044	0.64	0.67	0.66	0.71	0.65	0.72	0.61	0.71	0.52	0.75	0.68	0.89	0.46	0.56

Note: Empty cells indicate infeasible LP solutions; EDEA results after 100,000 generations.

Appendix B

Evolutionary Data Envelopment Analysis (EDEA)

```
/* EDEA Version 1.01                                                                    */
/* Evolutionary Data Envelopment Analysis                                               */
/* Decision Support System for the optimal selection of benchmarking partners.          */
/* EDEA proposes partners by following the concept of the input-oriented constant       */
/* returns of scale DEA model implemented by a genetic algorithm with floating point    */
/* representation (Michaelewicz GENOCOP system).                                        */
/* (c) 2000+ Karl Wöber, Vienna University for Economics and Business Administration    */
/* last modification: 2000-04-15                                                       */

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <time.h>
#include <math.h>
#define EPS 0.000001

double dea(unsigned int nx);
unsigned int check_output(unsigned int nx);
unsigned int check_input(unsigned int nx);
unsigned int check_feasibility(unsigned int nx);
void calc_boundaries(unsigned int nx);
void statistics(unsigned int gen);
void initreport(char *infile);
void report(unsigned int gen2);
void generation(unsigned int gen);
unsigned int select(void);
void new_to_old(void);
void scalepop(double umax, double uavg, double umin);

unsigned int init(void);
unsigned int crossover_sc(unsigned int x,unsigned int mate1, unsigned int mate2);
unsigned int crossover_sa(unsigned int x,unsigned int mate1, unsigned int mate2);
unsigned int crossover_wa(unsigned int x,unsigned int mate1, unsigned int mate2);
unsigned int mutation_um(unsigned int x,unsigned int mate);
unsigned int mutation_bm(unsigned int x,unsigned int mate);
```

```
unsigned int mutation_nm(unsigned int x,unsigned int mate);
unsigned int mutation_cm(unsigned int x,unsigned int mate);
unsigned int reproduction(unsigned int x,unsigned int mate);
double cooling(unsigned int t,double y);

unsigned int cases;                      /* Number of SBUs */
unsigned int oanz;                       /* Number of output constraints */
unsigned int iuanz;                      /* Number of uncontrollable input constraints */
unsigned int ianz;                       /* Number of input constraints */
unsigned int sbu;                        /* SBU case number */
double **rec;                            /* Data matrix cases x ianz+iuanz+oanz */
unsigned int x,y,z;                      /* General counter */
char datout[15]="";                      /* File name */
FILE *f1, *f2;                           /* I/O file handle */
double **oldpop;                         /* Old population */
double **newpop;                         /* New population */
double **newlb;                          /* New lower bounds defined by constraints */
double **newub;                          /* New upper bounds defined by constraints */
double **oldlb;                          /* Old lower bounds defined by constraints */
double **oldub;                          /* Old upper bounds defined by constraints */
double *lhsout;                          /* Left hand side vector */
double *bestsofar;                       /* Best result so far */
double maxfitness=0;                     /* Maximum fitness */
double minfitness=0;                     /* Minimum fitness */
double avgfitness=0;                     /* Average fitness */
double sumscaled=0;                      /* Total fitness scaled */
unsigned int bestchrom=0;                /* Best chromosom found so far */
const unsigned int popsize=40;           /* Population size */
const unsigned int maxgen=100000;        /* Maximum number of iterations */
double pm_um=.02;                        /* Uniform mutation probability */
double pm_bm=.01;                        /* Boundary mutation probability */
double pm_nm=.02;                        /* Non-uniform mutation probability */
double pm_sc=.1;                         /* Simple crossover probability */
double pm_sa=.1;                         /* Single arithmetic crossover probability */
double pm_wa=.1;                         /* Whole arithmetical crossover probability */
double pm_cm=.08;                        /* Conditional mutation probability */
unsigned int umc=0;                      /* Uniform mutation counter */
unsigned int bmc=0;                      /* Boundary mutation counter */
unsigned int nmc=0;                      /* Non-uniform mutation counter */
unsigned int ncm=0;                      /* Conditional mutation counter */
unsigned int scc=0;                      /* Simple crossover counter */
unsigned int sac=0;                      /* Single arithmetic crossover counter */
unsigned int wac=0;                      /* Whole arithmetical crossover counter */
unsigned int gen=0;                      /* Generation counter */
double pm_a=.25;                         /* Coefficient required for arithmetical crossover */
double pm_b=2;                           /* Coefficient required for non-uniform mutation */
unsigned int *anzmov;                    /* number of moveable genes */
unsigned int **moveable;                 /* array which holds the positions of the movable
genes */
unsigned int *crossable;                 /* array which holds the positions of the crossable
genes */

int main(int argc, char **argv)
{
  if (argc < 2) {printf("Usage: edea filename\n");}
  else if ((f1=fopen(argv[1],"r")) == NULL) {printf("Cannot open file.\n");}
  else {y=0; while (argv[1][y]!='.') {datout[y]=argv[1][y]; y++;} strcat(datout,".out");
      fscanf(f1, "%d %d %d %d\n", &cases, &oanz, &iuanz, &ianz, &sbu);
```

```
rec = (double **) malloc((iuanz+ianz+oanz)*sizeof(double *));
if (rec == NULL) {printf("Not enough memory\n"); return 0;}
for (y=0; y<iuanz+ianz+oanz; y++) {
  rec[y] = (double *) malloc(cases*sizeof(double));
  if (rec[y] == NULL) {printf("Not enough memory\n"); return 0;}}
for (x=0;x<cases;x++) {for (y=0;y<iuanz+ianz+oanz;y++) {
    fscanf(f1,"%lf", &rec[y][x]);}}
fclose(f1);
f2=fopen(datout,"w");

oldpop = (double **) malloc((popsize+1)*sizeof(double *));
if (oldpop == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  oldpop[x] = (double *) malloc((cases+3)*sizeof(double));
if (oldpop[x] == NULL) {printf("Not enough memory\n"); return 0;}}
newpop = (double **) malloc((popsize+1)*sizeof(double *));
if (newpop == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  newpop[x] = (double *) malloc((cases+3)*sizeof(double));
if (newpop[x] == NULL) {printf("Not enough memory\n"); return 0;}}

newlb = (double **) malloc((popsize+1)*sizeof(double *));
if (newlb == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  newlb[x] = (double *) malloc((cases)*sizeof(double));
if (newlb[x] == NULL) {printf("Not enough memory\n"); return 0;}}
newub = (double **) malloc((popsize+1)*sizeof(double *));
if (newub == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  newub[x] = (double *) malloc((cases)*sizeof(double));
if (newub[x] == NULL) {printf("Not enough memory\n"); return 0;}}

oldlb = (double **) malloc((popsize+1)*sizeof(double *));
if (oldlb == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  oldlb[x] = (double *) malloc((cases)*sizeof(double));
if (oldlb[x] == NULL) {printf("Not enough memory\n"); return 0;}}
oldub = (double **) malloc((popsize+1)*sizeof(double *));
if (oldub == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  oldub[x] = (double *) malloc((cases)*sizeof(double));
if (oldub[x] == NULL) {printf("Not enough memory\n"); return 0;}}

lhsout = (double *) malloc(cases*sizeof(double));
if (lhsout == NULL) {printf("Not enough memory\n"); return 0;}
bestsofar = (double *) malloc((cases+1)*sizeof(double));
if (bestsofar == NULL) {printf("Not enough memory\n"); return 0;}
anzmov = (unsigned int *) malloc(cases*sizeof(unsigned int));
if (anzmov == NULL) {printf("Not enough memory\n"); return 0;}
crossable = (unsigned int *) malloc(cases*sizeof(unsigned int));
if (crossable == NULL) {printf("Not enough memory\n"); return 0;}

moveable = (unsigned int **) malloc((popsize+1)*sizeof(unsigned int *));
if (moveable == NULL) {printf("Not enough memory\n"); return 0;}
for (x=0; x<(popsize+1); x++) {
  moveable[x] = (unsigned int *) malloc(cases*sizeof(unsigned int));
if (moveable[x] == NULL) {printf("Not enough memory\n"); return 0;}}
```

```
        srand(time(NULL));
        initreport(argv[1]);
        for (sbu=1;sbu<=cases;sbu++) {
                printf("EDEA for SBU %d ...\n",sbu);
                fprintf(f2,"EDEA for SBU %d ...\n",sbu);
                gen=0;

        if (init()) {
           statistics(gen);
           scalepop(maxfitness,avgfitness,minfitness);
           report(gen);
           new_to_old();

           /* Main EDEA */
           for (gen=1;gen<=maxgen;gen++) {
           generation(gen);
                   statistics(gen);
                   if (maxfitness-minfitness<.00001) {report(maxgen); break;}
                   scalepop(maxfitness,avgfitness,minfitness);
           report(gen);
           new_to_old();
          }
         }
       }
     fclose(f2);
   }
  return 0;
}


unsigned int init(void) {
   unsigned int i,imax;
   /* EDEA initialization */
   for (y=0;y<cases;y++) {bestsofar[y]=0;}
   bestsofar[cases]=999;
   printf("Initialize ",x);
   for (x=0;x<popsize;x++) {printf(".");
      if (x==0) {imax=5000000;} else {imax=10000;}
      for (i=0;i<imax;i++) {
           for (y=0;y<cases;y++) {newpop[x][y]=0;}
           while (1) {
                 y=rand()%(cases)+0; while (y==sbu-1) {y=rand()%(cases)+0;}
                 newpop[x][y]+=(rand()/(double)RAND_MAX);
                 if (check_output(x)) {break;}
           }
           if (check_input(x)) {
                 newpop[x][cases]=dea(x);
                 calc_boundaries(x);
                 break;
           }
        }
        if (i==imax) {
           if (x==0) {fprintf(f2,"Initialization for SBU %d failed.\n",sbu); return(0);}
           else {i=rand()%(x)+0;
              for (y=0;y<=cases;y++) {newpop[x][y]=newpop[i][y];}
              calc_boundaries(x);
           }
        }
```

```
        bestsofar[cases]=max(bestsofar[cases],newpop[x][cases]);
    }
    printf("\n");
    return(1);
}

void scalepop(double umax, double uavg, double umin) {
    unsigned int j;
    double a,b;
    const double fmultiple=4;
    /* Calculate scaling coefficients for linear scaling (Goldberg 1989) */
    if (umax==umin) {a=1; b=0;}
    else if (umin>(fmultiple*uavg-umax)/(fmultiple-1)) {
            a=(fmultiple-1)*uavg/(umax-uavg);
            b=uavg*(umax-fmultiple*uavg)/(umax-uavg);}
    else {
            a=uavg/(uavg-umin);
            b=-umin*uavg/(uavg-umin);
    }
    sumscaled=0;
    for (j=0;j<popsize;j++) {
            newpop[j][cases+2]=max(a*newpop[j][cases+1]+b,0);
            sumscaled+=newpop[j][cases+2];
    }
}


double dea(unsigned int nx) {
    unsigned int j,k;
    double lhs, emax=0;
    /* Calculate DEA efficiency scores */
    for (j=0;j<oanz+iuanz+ianz;j++) {
            lhs=0; for (k=0;k<cases;k++) {lhs+=rec[j][k]*newpop[nx][k];}
            if ((j>=oanz+iuanz) & (lhs/rec[j][sbu-1]>emax)) {emax=lhs/rec[j][sbu-1];}
    }
    return(emax);
}

unsigned int check_feasibility(unsigned int nx) {
    if (check_output(nx) & check_input(nx)) {return(1);} else {return(0);}
}

unsigned int check_output(unsigned int nx) {
    unsigned int j,k;
    /* Check output violation */
    double lhs;
    for (j=0;j<oanz;j++) {
            lhs=0; for (k=0;k<cases;k++) {lhs+=rec[j][k]*newpop[nx][k];}
            if (lhs-rec[j][sbu-1]<0) {return(0);}
    }
    return(1);
}

unsigned int check_input(unsigned int nx) {
    unsigned int j,k;
    /* Check input violation */
    double lhs;
    for (j=oanz;j<oanz+iuanz;j++) {
```

```
      lhs=0; for (k=0;k<cases;k++) {lhs+=rec[j][k]*newpop[nx][k];}
         if (rec[j][sbu-1]-lhs<0) {return(0);}
   }
   return(1);
}

void calc_boundaries(unsigned int nx) {
   /* Procedure calculates boundaries */
   unsigned int j,k;
   double v;
   anzmov[nx]=0;
   for (j=0;j<(oanz+iuanz);j++) {
      lhsout[j]=0; for (k=0;k<cases;k++) {lhsout[j]+=rec[j][k]*newpop[nx][k];}}
   for (k=0;k<cases;k++) {
         newlb[nx][k]=0; newub[nx][k]=999;
         for (j=0;j<oanz;j++) {
               v=(rec[j][k]*newpop[nx][k]-lhsout[j]+rec[j][sbu-1])/rec[j][k];
               if (v>newlb[nx][k]) {newlb[nx][k]=v;}
         }
         for (j=oanz;j<(oanz+iuanz);j++) {
               v=(rec[j][k]*newpop[nx][k]-lhsout[j]+rec[j][sbu-1])/rec[j][k];
               if (v<newub[nx][k]) {newub[nx][k]=v;}
         }
   }
   for (k=0;k<cases;k++) {
         if ((k!=sbu-1) & (newub[nx][k]-newlb[nx][k]>(EPS*2))) {
            moveable[nx][anzmov[nx]]=k; anzmov[nx]++;}
   }
}


void new_to_old(void) {
         unsigned int j,k;
         /* Procedure copies new to old_variables */
         for (j=0;j<popsize;j++) {
               for (k=0;k<cases;k++) {
                     oldpop[j][k]=newpop[j][k];
                     oldlb[j][k]=newlb[j][k];
               oldub[j][k]=newub[j][k];
               }
               oldpop[j][cases]=newpop[j][cases];          /* = efficiency scores */
               oldpop[j][cases+1]=newpop[j][cases+1];      /* = row fitness */
               oldpop[j][cases+2]=newpop[j][cases+2];      /* = scaled fitness */
         }
}

void statistics(unsigned int gen) {
      /* Calculate general statistics and perform cost to fitness transformation */
      unsigned int j;
      double sumfit;
      maxfitness=newpop[0][cases];
      for (j=1;j<popsize;j++) {maxfitness=max(maxfitness,newpop[j][cases]);}
      for (j=0;j<popsize;j++) {newpop[j][cases+1]=maxfitness*1.1-newpop[j][cases];}
      minfitness=newpop[0][cases+1];
      maxfitness=newpop[0][cases+1];
      sumfit=newpop[0][cases+1];
      bestchrom=0;
      for (j=1;j<popsize;j++) {
```

```
                sumfit+=newpop[j][cases+1];
                if (newpop[j][cases+1]>maxfitness) {
                   maxfitness=newpop[j][cases+1]; bestchrom=j;}
                if (newpop[j][cases+1]<minfitness) {
                   minfitness=newpop[j][cases+1];}
        }
        avgfitness=sumfit/popsize;
}

void report(unsigned int gen2) {
        /* Write output to file and screen */
        if (newpop[bestchrom][cases]<bestsofar[cases]) {
                for (x=0;x<=cases;x++) {bestsofar[x]=newpop[bestchrom][x];}
                printf("%4d SBU: %d Score: %5.3f BM:",gen,sbu,newpop[bestchrom][cases]);
                if (newpop[bestchrom][cases]<1) {for (x=0;x<cases;x++) {
                   if (bestsofar[x]>.05) {printf(" %d(%4.2f)",x+1,bestsofar[x]);};}}
                printf("\n");
        }

        if (gen2==maxgen) {
           printf("%4d SBU: %d Score: %5.3f BM:",gen,sbu,bestsofar[cases]);
           if (newpop[bestchrom][cases]<1) {for (x=0;x<cases;x++) {
              if (bestsofar[x]>.05) {printf(" %d(%4.2f)",x+1,bestsofar[x]);};}}
           printf("\n");
           fprintf(f2,"\nBest so far results for SBU %d after generation %d
                   (e=%f)\n",sbu,gen,bestsofar[cases]);
           for (x=0;x<cases;x++) {fprintf(f2,"SBU # %3d: %f\n",x+1,bestsofar[x]);}
        }
}


void initreport(char *infile) {
        /* Write startup parameters to output file */
        fprintf(f2,"Inputfile: %s\n\n", infile);
        fprintf(f2,"Cases %d\nInput (c) %d\nInput (u) %d\nOutput %d\nSBU %d\n\n",
              cases,ianz,iuanz,oanz,sbu);
        fprintf(f2,"Population size %d\nMaximum # of generations %d\n\nProbabilities:\n",
              popsize,maxgen);
        fprintf(f2,"- Uniform mutation %f\n- Boundary mutation %f\n- Non-uniform mutation
              %f\n- Conditional mutation %f\n- Simple crossover %f\n- Single arithmetical
              crossover %f\n- Whole arithmetical crossover %f\n\n",
              pm_um,pm_bm,pm_nm,pm_cm,pm_sc,pm_sa,pm_wa);
}

void generation(unsigned int gen) {
   /* Generates a new population through select, crossover, and mutation */
   /* GA operators return increment in population size */
   unsigned int popincr,j,k;
   double rnumber;
   x=0;
   while (x<popsize) {j=select();
      rnumber=rand()/(double)RAND_MAX;
      if (rnumber>pm_wa+pm_sa+pm_sc+pm_cm+pm_um+pm_bm+pm_nm)
         {popincr=reproduction(x,j);}
      else if (rnumber>pm_wa+pm_sa+pm_sc+pm_cm+pm_um+pm_bm) {popincr=mutation_nm(x,j);}
      else if (rnumber>pm_wa+pm_sa+pm_sc+pm_cm+pm_um)        {popincr=mutation_bm(x,j);}
      else if (rnumber>pm_wa+pm_sa+pm_sc+pm_cm)              {popincr=mutation_um(x,j);}
      else if (rnumber>pm_wa+pm_sa+pm_sc)                    {popincr=mutation_cm(x,j);}
```

```
        else {
            k=select(); while(k==j) {k=select();}
      if            (rnumber>pm_wa+pm_sa)            popincr=crossover_sc(x,j,k);}
      else if       (rnumber>pm_wa)                  {popincr=crossover_sa(x,j,k);}
      else                                           {popincr=crossover_wa(x,j,k);}
        }
        x+=popincr;
   }
   for (j=0;j<popsize;j++) {
          newpop[j][cases]=dea(j);
          calc_boundaries(j);
   }
}

unsigned int select(void) {
   /* Roulette wheel selection */
   double rnumber,partsum;
   unsigned int j;
   rnumber=(rand()/(double)RAND_MAX)*sumscaled;
   partsum=0;
   for (j=0;j<popsize-1;j++) {
          partsum+=oldpop[j][cases+2];
          if (partsum>rnumber) {break;}
   }
   return(j);
}


unsigned int crossover_sc(unsigned int x, unsigned int mate1, unsigned int mate2) {
   /* Simple crossover (Michalewicz 1996) */
   unsigned int jcross,j;
   double a=1;
   scc++;
   jcross=rand()%(cases)+0;
   for (j=0;j<jcross;j++) {newpop[x][j]=oldpop[mate1][j];
      newpop[x+1][j]=oldpop[mate2][j];}
   if (jcross<cases) {while (a>=0) {
        for (j=jcross;j<cases;j++) {
          newpop[x][j]=oldpop[mate2][j]*a+oldpop[mate1][j]*(1-a);
          newpop[x+1][j]=oldpop[mate1][j]*a+oldpop[mate2][j]*(1-a);
          }
        if (check_feasibility(x)+check_feasibility(x+1)==2) {break;} else {a-=.2;}
     }
   }
   return(2);
}

unsigned int crossover_wa(unsigned int x, unsigned int mate1, unsigned int mate2) {
   /* Whole arithmetical crossover (Michalewicz 1996) */
   unsigned int j;
   wac++;
   for (j=0;j<cases;j++) {
          newpop[x][j]=oldpop[mate1][j]*pm_a+oldpop[mate2][j]*(1-pm_a);
          newpop[x+1][j]=oldpop[mate2][j]*pm_a+oldpop[mate1][j]*(1-pm_a);
   }
   return(2);
}
```

```
unsigned int crossover_sa(unsigned int x, unsigned int mate1, unsigned int mate2) {
   /* Single arithmetical crossover (Michalewicz 1996) */
   /* With optimization of gene selection */
   unsigned int jcross,j;
   double cincr,alpha,beta,gamma,delta;
   unsigned int anzcross=0;
   for (j=0;j<cases;j++) {if ((oldpop[mate1][j]-oldpop[mate2][j]>EPS*2) |
      (oldpop[mate2][j]-oldpop[mate1][j]>EPS*2)) {crossable[anzcross]=j; anzcross++;}}
   if (anzcross==0) {j=reproduction(x,mate1); j=reproduction(x+1,mate2); return(2);}
   sac++; jcross=crossable[(rand()%(anzcross)+0)];
   alpha=(oldlb[mate2][jcross]-oldpop[mate2][jcross])
            /(oldpop[mate1][jcross]-oldpop[mate2][jcross]);
   beta =(min(oldub[mate1][jcross],1+oldlb[mate1][jcross])-oldpop[mate1][jcross])
            /(oldpop[mate2][jcross]-oldpop[mate1][jcross]);
   gamma=(oldlb[mate1][jcross]-oldpop[mate1][jcross])
            /(oldpop[mate2][jcross]-oldpop[mate1][jcross]);
   delta=(min(oldub[mate2][jcross],1+oldlb[mate2][jcross])-oldpop[mate2][jcross])
            /(oldpop[mate1][jcross]-oldpop[mate2][jcross]);
   for (j=0;j<cases;j++) {if (j==jcross) {if (oldpop[mate1][j]>oldpop[mate2][j]){
      cincr=max(alpha,beta)+(EPS+(rand()/(double)RAND_MAX)*(double)(1-EPS*2))
            *(min(gamma,delta)-max(alpha,beta));} else {
      cincr=max(gamma,delta)+(EPS+(rand()/(double)RAND_MAX)*(double)(1-EPS*2))
            *(min(alpha,beta)-max(gamma,delta));}
               newpop[x][j]=oldpop[mate2][j]*cincr+oldpop[mate1][j]*(1-cincr);
               newpop[x+1][j]=oldpop[mate1][j]*cincr+oldpop[mate2][j]*(1-cincr);
         } else {
            newpop[x][j]=oldpop[mate1][j];
            newpop[x+1][j]=oldpop[mate2][j];
         }
   }
   return(2);
}

unsigned int reproduction(unsigned int x, unsigned int mate) {
   unsigned int j;
   for (j=0;j<cases;j++) {newpop[x][j]=oldpop[mate][j];}
   return(1);
}

unsigned int mutation_um(unsigned int x, unsigned int mate) {
   /* Uniform mutation (Michalewicz 1996) */
   /* With optimization of gene selection */
   unsigned int j,k;
   if (anzmov[mate]==0) {j=reproduction(x,mate); return(1);}
   k=moveable[mate][(rand()%(anzmov[mate])+0)];
   umc++;
   for (j=0;j<cases;j++) {
         if (j==k) {newpop[x][j]=oldlb[mate][j]+EPS+(rand()/(double)RAND_MAX)
                  *(min(oldub[mate][j],1+oldlb[mate][j])-oldlb[mate][j]-2*EPS);
         } else {newpop[x][j]=oldpop[mate][j];}
   }
   return(1);
}

unsigned int mutation_bm(unsigned int x, unsigned int mate) {
   /* Boundary mutation (Michalewicz 1996) */
   /* With optimization of gene selection */
   unsigned int j,k;
```

```
   if (anzmov[mate]==0) {j=reproduction(x,mate); return(1);}
   k=moveable[mate][(rand()%(anzmov[mate])+0)];
   bmc++;
   for (j=0;j<cases;j++) {
            if (j==k) {if ((rand()/(double)RAND_MAX)<.5) {newpop[x][j]=oldlb[mate][j]+EPS;
            } else {newpop[x][j]=min(oldlb[mate][j]+1,oldub[mate][j]-EPS);}}
            else {newpop[x][j]=oldpop[mate][j];}
   }
   return(1);
}

unsigned int mutation_nm(unsigned int x, unsigned int mate) {
   /* Non-uniform mutation (Michalewicz 1996) */
   /* With optimization of gene selection */
   unsigned int j,k;
   if (anzmov[mate]==0) {j=reproduction(x,mate); return(1);}
   k=moveable[mate][(rand()%(anzmov[mate])+0)];
   nmc++;
   for (j=0;j<cases;j++) {
            if (j==k) {
               if (rand()%(2)+0) {newpop[x][j]=oldpop[mate][j]
                                 -cooling(gen,oldpop[mate][j]-oldlb[mate][j]-EPS);
               } else {newpop[x][j]=oldpop[mate][j]
                 +cooling(gen,min(oldub[mate][j],1+oldlb[mate][j])-oldpop[mate][j]-EPS);}
            } else {newpop[x][j]=oldpop[mate][j];}
   }
   return(1);
}

double cooling(unsigned int t, double y) {
      /* Simulated annealing algorithm */
         return(y*(1-pow((rand()/(double)RAND_MAX),pow(1-(double)t/maxgen,pm_b))));
}


unsigned int mutation_cm(unsigned int x, unsigned int mate) {
   /* Conditional mutation */
   /* With optimization of gene selection */
   unsigned int j,k,l,anzcross=0;
   double r,a=1;
   for (j=0;j<cases;j++) {if ((oldpop[mate][j]>0) & (j!=sbu-1)) {
        crossable[anzcross]=j; anzcross++;}}
   if (anzcross==0) {j=reproduction(x,mate); return(1);}
   ncm++;
   l=crossable[(rand()%(anzcross)+0)];
   k=(rand()%(cases)+0); while ((k==l) | (k==sbu-1)) {k=(rand()%(cases)+0);}
   while (a>=0) {
         for (j=0;j<cases;j++) {newpop[x][j]=oldpop[mate][j];}
         r=(rand()/(double)RAND_MAX);
         newpop[x][l]=oldpop[mate][l]-r*a*oldpop[mate][l];
         newpop[x][k]=oldpop[mate][k]+r*a*oldpop[mate][l];
         if (check_feasibility(x)) {break;} else {a-=.2;}
   }
   return(1);
}
```

References

Adams, D.J. (1991) Do corporate failure prediction models work? *International Journal of Contemporary Hospitality Management* 3(4), 6–13.

Adams, D.J. (1995) Methods of predicting financial failure in the hotel industry. In: Harris, P.J. (ed.) *Accounting and Finance for the International Hospitality Industry*. Butterworth-Heinemann, Oxford, pp. 19–38.

Adams, D.J. and Kwansa, F. (1992) Corporate performance appraisal in the international hospitality industry. In: Teare, R. and Olsen, M. (eds) *International Hospitality Management: Corporate Strategy in Practice*. Pitman, London, pp. 34–49.

Adler, N. and Berechman, J. (2001) Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis. *Transport Policy* 8(3), 171–181.

Ahmed, P.K. and Rafiq, M. (1998) Integrated benchmarking: a holistic examination of select techniques for benchmarking analysis. *Benchmarking for Quality Management and Technology* 5(3), 225–242.

Ahn, T. and Seiford, L. (1990) Sensitivity of DEA to models and variable sets in a hypothesis-test setting: the efficiency of university operations. In: Ijiri, Y. (ed.) *Creative and Innovative Approaches to the Science of Management*. Quorum, Westport, Connecticut, pp. 191–208.

Aigner, D.J., Lovell, C.A.K. and Schmidt, P.J. (1977) Formulation and estimation of stochastic frontier production function model. *Journal of Econometrics* 6(1), 21–37.

Aitkin, M. and Longford, N. (1986) Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society – Series A* 149, 419–461.

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N. and Csaki, F. (eds) *Proceedings of the Second International Symposium on Information Theory*. Akademial Kiado, Budapest, pp. 267–281.

Alavi, J. and Yasin, M.M. (2000) A systematic approach to tourism policy. *Journal of Business Research* 48(2), 147–157.

Ali, A.I. (1992) Streamlined computation for data envelopment analysis. *European Journal of Operational Research* 64(1), 61–67.

Ali, A.I. (1994) Computational aspects of DEA. In: Charnes, A.C., Cooper, W.W., Lewin, A.Y. and Seiford, L.M. (eds) *Data Envelopment Analysis: Theory, Methodology and Application*. Kluwer, Boston, Massachusetts, pp. 63–88.

Ali, A.I. and Seiford, L.M. (1993) The mathematical programming approach to efficiency measurement. In: Fried, H., Lovell, C.A.K. and Schmidt, S. (eds) *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press, London, pp. 120–159.

Al-Shammari, M. (1999) Optimisation modelling for estimating and enhancing relative efficiency with application to industrial companies. *European Journal of Operational Research* 115(3), 488–496.

Altman, E.I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23(4), 589–609.

Altman, E.I., Haldeman, R. and Narayanan, P. (1977) ZETA analysis: a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* 1, 29–54.

American Hotel & Motel Association (1996) *Uniform System of Accounts for the Lodging Industry*, 9th revised edn. Educational Institute, New York.

Andersen, P. and Petersen, N.C. (1993) A procedure for ranking efficient units in data envelopment analysis. *Management Science* 39(10), 1261–1264.

Anderson, D.R., Sweeney, D.J. and Williams, T.A. (1997) *An Introduction to Management Science. Quantitative Approaches to Decision Making*, 8th edn. West Publishing, St Paul, Minnesota.

Anderson, R.I., Fish, M., Xia, Y. and Michello, F. (1999) Measuring efficiency in the hotel industry: a stochastic frontier approach. *Hospitality Management* 18(1), 45–57.

Anderson, R.I., Fok, R. and Scott, J. (2000) Hotel industry efficiency: an advanced linear programming examination. *American Business Review* (January), 40–48.

Andersson, T.D. (1996) Traditional key ratio analysis versus data envelopment analysis: a comparison of various measurements of productivity and efficiency in restaurants. In: Johns, N. (ed.) *Productivity Management in Hospitality and Tourism*. Cassell, London, pp. 209–226.

Athanassopoulos, A.D. and Curram, S.P. (1996) A comparison of data envelopment analysis and artificial neural networks as tools for assessing the efficiency of decision making units. *Journal of the Operational Research Society* 47, 1000–1016.

Athanassopoulos, A.D., Lambroukos, N. and Seiford, L. (1999) Data envelopment scenario analysis for setting targets to electricity generating plants. *European Journal of Operational Research* 115(3), 413–428.

Atkinson, H. and Brown, J.B. (2001) Rethinking performance measures: assessing progress in UK hotels. *International Journal of Contemporary Hospitality Management* 13(3), 128–135.

Atkinson, H., Berry, A. and Jarvis, R. (1995) *Business Accounting for Hospitality and Tourism*. Chapman & Hall, London.

Avkiran, N.K. (1999) *Productivity Analysis in the Services Sector with Data Envelopment Analysis*. Avkiran Publ., Camira, Queensland.

Back, B., Irjala, M., Sere, K. and Vanharanta, H. (1995) Competitive financial benchmarking using self-organising maps. In: Vasarhelyi, M.A. (ed.) *Artificial Intelligence in Accounting and Auditing: the Use of Expert Systems*. Markus Wiener Publishers, New York, pp. 69–81.

Bäck, T. and Schwefel, H.-P. (1993) An overview of evolutionary algorithms for parameter optimisation. *Evolutionary Computation* 1(1), 1–23.

Baker, M. and Riley, M. (1994) New perspectives on the productivity of hotels: some advances and new directions. *International Journal of Hospitality Management* 13(4), 297–311.

Baker, M.J., Black, C.D. and Hart, S. (1988) The competitiveness of British industry: what really makes the difference? *European Journal of Marketing* 22(2), 70–85.

Ballestero, E. (1999) Measuring efficiency by a single price system. *European Journal of Operational Research* 115(3), 616–623.

Banker, R.D. (1984) Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research* 17(1), 35–44.

Banker, R.D. (1989) Econometric estimation and data envelopment analysis. *Research in Governmental and Non-profit Accounting* 5, 231–243.

Banker, R.D. (1993) Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. *Management Science* 39(10), 1265–1273.

Banker, R.D. and Morey, R.C. (1986a) The use of categorial variables in data envelopment analysis. *Management Science* 32(12), 1613–1627.

Banker, R.D. and Morey, R.C. (1986b) Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research* 34(4), 513–521.

Banker, R.D. and Morey, R.C. (1989) Incorporating value judgements in efficiency analysis. *Research in Governmental and Non-profit Accounting* 5, 245–267.

Banker, R.D. and Thrall, R.M. (1992) Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research* 62(1), 74–84.

Banker, R.D., Charnes, A.C. and Cooper, W.W. (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30(9), 1078–1092.

Banker, R.D., Das, S. and Datar, S. (1989) Analysis of cost variances for management control in hospitals. *Research in Governmental and Non-profit Accounting* 5, 269–291.

Banker, R.D., Gadh, V.M. and Gorr, W.L. (1993) A Monte Carlo comparison of two production frontier estimation methods: corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research* 67(3), 332–343.

Banker, R.D., Chang, H. and Cooper, W.W. (1996) Equivalence and implementation of alternative methods for determining returns to scale in data envelopment analysis. *European Journal of Operational Research* 89(3), 473–481.

Banker, R.D., Chang, H. and Das, S. (1998) Standard estimation, standard tightness, and benchmarking: a method with an application to nursing services. *Journal of Management Accounting Research* 10, 133–153.

Banker, R.D., Potter, G. and Srinivasan, D. (2000) An empirical investigation of an incentive plan based non-financial performance measures. *The Accounting Review* 75(1), 65–92.

Banzhaf, W., Nordin, P., Keller, R.E. and Francone, F.D. (1998) *Genetic Programming. An Introduction on the Automatic Evolution of Computer Programs and its Applications*. Morgan Kaufmann, San Francisco, California.

Barnes, P. (1982) Methodological implications of non-normally distributed financial ratios. *Journal of Business Finance and Accounting* 9(1), 51–62.

Barr, R.S. and Durchholz, M.L. (1997) Parallel and hierarchical decomposition approaches for solving large-scale data envelopment analysis models. *Annals of Operations Research* 73, 339–372.

Barr, R.S. and Siems, T.F. (1997) Bank failure prediction using DEA to measure management quality. In: Barr, R.S. (ed.) *Interfaces in Computer Science and Operations Research*. Kluwer Academic, Boston, Massachusetts, pp. 341–366.

Barr, R.S., Seiford, L.M. and Siems, T.F. (1993) An envelopment-analysis approach to measuring the managerial quality of banks. *Annals of Operations Research* 45, 1–19.

Barr, R.S., Seiford, L.M. and Siems, T.F. (1994) Forecasting bank failure: a non-parametric frontier estimation approach. *Reserches Economiques de Lovain* 60(4), 417–429.

Barrow, M. and Wagstaff, A. (1989) Efficiency measurement in the public sector: an appraisal. *Fiscal Studies* 10(1), 72–97.

Bauer, P.W. (1990) Recent developments in econometric estimation of frontiers. *Journal of Econometrics* 46(1/2), 39–56.

Beasley, D., Bull, D.R. and Martin, R.R. (1993) A sequential niche technique for multimodal function optimisation. *Evolutionary Computation* 1(2), 101–125.

Bell, R.A. and Morey, R.C. (1994) The search for appropriate benchmarking partners: a macro approach and application to corporate travel management. *Omega, International Journal of Management Science* 22(5), 477–490.

Bell, R.A. and Morey, R.C. (1995) Increasing the efficiency of corporate travel management through macro benchmarking. *Journal of Travel Research* 33(3), 11–21.

Berger, A.N., Hunter, W.C. and Timme, S.G. (1993) The efficiency of financial institutions: a review and preview of research past, present, and future. *Journal of Banking and Finance* 17 (April), 223–227.

Bergin, S., Jago, L.K. and Deery, M. (2000) Benchmarking in the hospitality industry: an important but misinterpreted concept. *Australian Journal of Hospitality Management* 7 (2), 25–38.

Bernstein, L.A. (1978) *Financial Statement Analysis*. Richard D. Irwin, Homewood, Illinois.

Bertels, K., Jaques, J.M., Neuberg, L. and Gatot, L. (1999) Qualitative company performance evaluation: linear discriminant analysis and neural network models. *European Journal of Operational Research* 115(3), 608–615.

Bifulco, R. and Bretschneider, S. (2001) Estimating school efficiency. A comparison of methods using simulated data. *Economics of Education Review* 20(4), 417–429.

Bjurek, H., Hjalmarsson, L. and Førsund, F.R. (1990) Deterministic parametric and nonparametric estimation of efficiency in service production. A comparison. *Journal of Econometrics* 46(1/2), 213–227.

Blum, M. (1974) Failing company discriminant analysis. *Journal of Accounting Research* 12(Spring), 1–25.

Bogan, C.E. and English, M.J. (1994) *Benchmarking for Best Practices: Winning Through Innovative Adaptation*. McGraw-Hill, New York.

Boger, C.A., Cai, L.A. and Lin, Li-C. (1999) Benchmarking: comparing discounted business rates among lodging companies. *Journal of Hospitality and Tourism Research* 23(3), 256–267.

Boljunčić, V. (1999) A note on robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research* 112(1), 240–244.

Bonoma, T.V. (1989) Marketing performance – what do you expect? *Harvard Business Review* 89(5), 44–47.

Booker, L. (1987) Improving search in genetic algorithms. In: Davis, L. (ed.) *Genetic Algorithms and Simulated Annealing*. Pitman, London, pp. 61–73.

Bottomley, V. (1995) *Tourism: Competing With the Best. Benchmarking for Smaller Hotels.* KPMG study, Department of National Heritage, London.

Box, G.E.P. (1957) Evolutionary operation: a method of increasing industrial productivity. *Journal of the Royal Statistical Society – Series C* 6(2), 81–101.

Bozdogan, H. (1994) Mixture model cluster analysis using model selection criteria and a new informational measure of complexity. In: Bozdogan, H. (ed.) *Multivariate Statistical Modelling.* Kluwer Academic, Dordrecht, pp. 69–113.

Breiter, D. and Kline, S.F. (1995) Benchmarking quality management in hotels. *FIU Hospitality Review* 13(2), 45–60.

Brignall, S. and Ballantine, J. (1996) Performance measurement in service businesses revisited. *International Journal of Service Industry Management* 7(1), 6–31.

Brockett, P.L. and Golany, B. (1996) Using rank statistics for determining programmatic efficiency differences in data envelopment analysis. *Management Science* 42(3), 466–472.

Brander-Brown, J.B. and McDonnell, B. (1995) The balanced-score card: short-term guest or long-term resident? *International Journal of Contemporary Hospitality Management* 7(2/3), 7–11.

Buckley, P.J., Pass, C.L. and Prescott, K. (1988) Measures of international competitiveness: a critical survey. *Journal of Marketing Management* 4(2), 175–200.

Buzzell, R.D., Gale, B.T. and Sultan, R.G.M. (1975) Market share – a key to profitability. *Harvard Business Review* 53(1), 97–106.

Camp, R.C. (1989) *Benchmarking: the Search for Industry Best Practices that Lead to Superior Performance.* American Society for Quality Control Quality Press, Milwaukee, Wisconsin.

Camp, R.C. (1995) *Business Process Benchmarking: Finding and Implementing Best Practices.* American Society for Quality Control Quality Press, Milwaukee, Wisconsin.

Canestrelli, E. and Costa, P. (1991) Tourist carrying capacity: a fuzzy approach. *Annals of Tourism Research* 18(2), 295–311.

Caporaletti, L.E., Dulá, J.H. and Womer, N.K. (1999) Performance evaluation based on multiple attributes with nonparametric frontiers. *Omega, International Journal of Management Science* 27(6), 637–645.

Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982) The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica* 5(6), 1393–1414.

Chandra, P., Cooper, W.W., Li, S. and Rahman, A. (1998) Using DEA to evaluate 29 Canadian textile companies – considering returns to scale. *International Journal of Production Economics* 54(2), 129–141.

Charnes, A.C. and Cooper, W.W. (1985) Preface to topics in data envelopment analysis. *Annals of Operations Research* 2, 59–94.

Charnes, A.C., Cooper, W.W. and Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2(6), 429–444.

Charnes, A.C., Cooper, W.W. and Rhodes, E. (1981) Evaluating program and managerial efficiency: an application of the data envelopment analysis to program follow through. *Management Science* 27(6), 668–697.

Charnes, A.C., Clark, T., Cooper, W.W. and Golany, B. (1985a) A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the U.S. Air Force. *Annals of Operations Research* 2, 95–112.

Charnes, A.C., Cooper, W.W., Lewin, A.Y., Morey, R.C. and Rousseau, J. (1985b) Sensitivity and stability analysis in DEA. *Annals of Operations Research* 2, 139–156.

Charnes, A.C., Cooper, W.W., Wie, Q.L. and Huang, Z.M. (1989) Cone ratio data envelopment analysis and multi-objective programming. *International Journal of Systems Science* 20(7), 1099–1118.

Charnes, A.C., Cooper, W.W., Huang, Z. and Sun, D. (1990) Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *Journal of Econometrics* 46(1/2), 73–91.

Charnes, A.C., Rousseau, J. and Semple, J. (1993) An effective non-Archimedan anti-degeneracy/cycling linear programming method especially for data envelopment analysis and like models. *Annals of Operations Research* 46/47, 271–278.

Charnes, A.C., Cooper, W.W., Golany, B., Learner, D.B., Phillips, F.Y. and Rousseau, J.J. (1994a) A multiperiod analysis of market segments and brand efficiency in the competitive carbonated beverage industry. In: Charnes, A.C., Cooper, W.W., Lewin, A.Y. and Seiford, L.M. (eds) *Data Envelopment Analysis: Theory, Methodology, and Application*. Kluwer, Boston, Massachusetts, pp. 145–165.

Charnes, A.C., Cooper, W.W., Lewin, A.Y. and Seiford, L.M. (1994b) *Data Envelopment Analysis: Theory, Methodology and Applications*. Kluwer, Boston, Massachusetts.

Cheng, F.Y. and Li, D. (1997) Genetic algorithm development for multiobjective optimisation of structures. *Journal of Structural Engineering* 123(9), 1252–1261.

Chin, J., Barney, W. and O'Sullivan, H. (1995) Best accounting practice in hotels: a guide for other industries? *Management Accounting* 73(11), 57–59.

Codling, S. (1998) Benchmarking. In: Lock, D. (ed.) *The Gower Handbook of Management*, 4th edn. Gower, Aldershot, UK, pp. 182–198.

Coelli, T.J. (1996) *A Guide to DEAP Version 2.1: a Data Envelopment Analysis (Computer) Program*. CEPA Working Papers No. 8/96, University of New England, Armidale, Australia.

Coelli, T.J. (1997) *A Multi-Stage Methodology for the Solution of Orientated DEA Models*. Mimeo, Centre for Efficiency and Productivity Analysis, University of New England, Armidale, Australia.

Coelli, T., Rao, D.S.P. and Battese, G.E. (1998) *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic, Boston, Massachusetts.

Collier, P. and Gregory, A. (1995) The practice of management accounting in hotel groups. In: Harris, P.J. (ed.) *Accounting and Finance for the International Hotel Industry*. Butterworth-Heinemann, Oxford, pp. 137–160.

Coltman, M.M. (1998) *Hospitality Management Accounting*. CBI Publishing, Boston, Massachusetts.

Cook, S. (1995) *Practical Benchmarking: a Manager's Guide to Creating a Competitive Advantage*. Kogan Page, London.

Cook, W.D., Kress, M. and Seiford, L.M. (1993) On the use of ordinal data in envelopment analysis. *Journal of the Operational Research Society* 44(2), 133–140.

Cooper, R.G., Park, K.S. and Yu, G. (1999) IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Management Sciences* 45(4), 597–607.

Cooper, W.W., Thompson, R.G. and Thrall, R.M. (eds) (1996) Extensions and new developments in data envelopment analysis. *Annals of Operations Research* 66.

Cooper, W.W., Seiford, L.M. and Tone, K. (2000) *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic, Boston, Massachusetts.

Cornwell, C., Schmidt, P. and Sickles, R.C. (1990) Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics* 46(1/2), 185–200.

Cross, M. (1998) *The Benchmarking Sourcebook*. Batsford, London.

Cubbin, J. and Tzanidakis, G. (1998) Techniques for analysing company performance. *Business Strategy Review* 9(4), 37–46.

Czarnecki, M.T. (1999) *Managing by Measuring: How to Improve Your Organizations's Performance Through Effective Benchmarking*. Amacom, New York.

Dantzig, G.B. (1963) *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey.

Dantzig, G.B. and Thapa, M.N. (1997) *Linear Programming*. Springer, New York.

Darwin, C. (1859) *The Origin of Species*. John Murray, London.

Darwin, C. (1985) *The Origin of Species*. Penguin Classics, London.

Davis, L. (1991) *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York.

Dawid, H. (1999) *Adaptive Learning by Genetic Algorithms. Analytical Results and Applications to Economic Models*, 2nd edn. Springer, Berlin.

Day, G.S. and Wensley, R. (1988) Assessing advantage: a framework for diagnosing competitive superiority. *Journal of Marketing* 52(2), 1–20.

Deakin, E.B. (1976a) Distributions of financial accounting ratios: some empirical evidence. *The Accounting Review* 51, 90–96.

Deakin, E.B. (1976b) A discriminant analysis of predictors of business failure. *Journal of Accounting Research* 14, 167–179.

Deb, K. (1996) Genetic algorithms for function optimisation. In: Herrera, F. and Verdegay, J.L. (eds) *Studies in Fuzziness and Soft Computing: Genetic Algorithms and Soft Computing*. Physica, Heidelberg, pp. 3–29.

De Jong, K.A. (1975) An Analysis of Behavior of a Class of Genetic Adaptive Systems. PhD thesis, Department of Computer and Communication Sciences, University of Michigan, Ann Arbor, Michigan.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society – Series B* 39(1), 1–38.

Denton, G.A. and White, B. (2000) Implementing a balanced-scorecard approach to managing hotel operations: the case of White Lodging Services. *Cornell Hotel and Restaurant Administration Quarterly* 41(1), 94–108.

DeSarbo, W.S. and Cron, W.L. (1988) A maximum likelihood methodology for clusterwise linear regression. *Journal of Classification* 5(2), 249–282.

Dielman, T.E. (1989) *Pooled Cross-Sectional and Time-Series Data Analysis*. Marcel Dekker, New York.

Dillon, W. and Kumar, A. (1994) Latent structure and other mixture models in marketing: an integrative survey and overview. In: Bagozzi, R.P. (ed.) *Advanced Methods in Marketing Research*. Blackwell, Cambridge, Massachusetts, pp. 295–351.

Dorsey, R.E. and Mayer, W.J. (1995) Genetic algorithms for estimating problems with multiple optima, nondifferentiability, and other irregular features. *Journal of Business and Economic Statistics* 13(1), 53–66.

Downie, N. (1995) The use of accounting information in hotel marketing decisions. In: Harris, P.J. (ed.) *Accounting and Finance for the International Hotel Industry*. Butterworth-Heinemann, Oxford, pp. 202–222.

Downie, N. (1997) The use of accounting information in hotel marketing decisions. *International Journal of Hospitality Management* 16(3), 305–312.

Doyle, J.R. and Green, R.H. (1994) Strategic choice and data envelopment analysis: comparing computers across many attributes. *Journal of Information Technology* 9, 61–69.

Doyle, J.R., Arthurs, A.J., Green, R.H., McAulay, L., Pitt, M.R., Bottomley, P.A. and Evans, W. (1996) The judge, the model of the judge, and the model of the judged as judge: analyses of the UK 1992 research assessment exercise data for business and management studies. *Omega, International Journal of Management Science* 24(1), 13–28.

Dubé, L., Enz, C.A., Renaghan, L.M. and Siguaw, J.A. (1999) *American Lodging Excellence: the Key to Best Practices in the U.S. Lodging Industry*. American Express and the American Hotel Foundation, Washington, DC.

Durocher, J.F. and Niman, N.B. (1993) Information technology: management effectiveness and guest services. *Hospitality Research Journal* 17(1), 121–131.

Epstein, M.K. and Henderson, J.C. (1989) Data envelopment analysis for managerial control and diagnostics. *Decision Sciences* 20(1), 90–119.

Eshelman, L. and Schaffer, J.D. (1993) Real-coded genetic algorithms and interval-schemata. *Foundations of Genetic Algorithm II*. Morgan Kaufmann, San Mateo, California.

Everitt, B.S. and Hand, D.J. (1981) *Finite Mixture Distributions*. Chapman and Hall, London.

Färe, R. and Grosskopf, S. (1996) *Intertemporal Production Frontiers: With Dynamic DEA*. Kluwer Academic Publishers, Boston, Massachusetts.

Färe, R., Grosskopf, S. and Logan, J. (1987) The comparative efficiency of Western coal-fired steam-electric generating plants: 1977–1979. *Engineering Costs and Production Economics* 11, 21–30.

Färe, R., Grosskopf, S., Lindgren, B. and Roos, P. (1992) Productivity changes in Swedish pharmacies 1980–1989: a non-parametric approach. *Journal of Productivity Analysis* 3(1/2), 85–101.

Färe, R., Grosskopf, S. and Lovell, C.A.K. (1994) *Production Functions*. Cambridge University Press, Cambridge.

Färe, R. and Lovell, C.A.K. (1978) Measuring the technical efficiency of production. *Journal of Economic Theory* 19(1), 150–162.

Farrell, M.J. (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society – Series A* 3, 253–290.

Fay, C.T., Rhoads, R.C. and Rosenblatt, R.L. (1976) *Managerial Accounting for Hospitality Service Industries*, 2nd edn. William C. Brown Publishers, Dubuque, Iowa.

Fenton, L.S., Fowler, N.A. and Parkinson, G.S. (1989) *Hotel Accounts and Their Audit*, 2nd edn. Institute of Chartered Accountants in England and Wales, London.

Ferrier, G.D. and Lovell, C.A.K. (1990) Measuring cost efficiency in banking: econometric and linear programming evidence. *Journal of Econometrics* 46(1/2), 229–245.

Fieldsend, S., Longford, N. and McLeay, S. (1987) Industry effects and the proportionality assumption in ratio analysis: a variance component analysis. *Journal of Business, Finance and Accounting* 14(4), 497–517.

Fisher, J.G. (1995) Use of nonfinancial performance measures. In: Young, S.M. (ed.) *Management Accounting*. Prentice Hall, Englewood Cliffs, New Jersey, pp. 329–335.

Fitzgerald, L., Johnston, R., Brignall, T.J., Silvestro, R. and Voss, C. (1991) *Performance Measurement in Service Businesses*. Chartered Institute of Management Accountants, London.

Fonseca, C.M. and Fleming, P.J. (1995) An overview of evolutionary algorithms in multiobjective optimisation. *Evolutionary Computation* 3(1), 1–16.

Førsund, F.R., Lovell, C.A.K. and Schmidt, P. (1980) A survey of frontier production functions and of their relationship to efficiency measurement. *Journal of Econometrics* 13(1), 5–25.

Frazier, G.L. and Howell, R.D. (1987) Business definition and performance. *Journal of Marketing* 47(2), 59–67.

Ganley, J.A. and Cubbin, J.S. (1992) *Public Sector Efficiency Measurement. Applications of Data Envelopment Analysis*. Elsevier, Amsterdam.

Garvin, D.A. (1993) Building a learning organisation. *Harvard Business Review* 71(4), 78–91.

Geller, A.N. (1985a) Tracking the critical success factors for hotel companies. *Cornell Hotel & Restaurant Administration Quarterly* 25(2), 76–81.

Geller, A.N. (1985b) The current state of hotel information systems. *Cornell Hotel & Restaurant Administration Quarterly* 25(2), 14–17.

Gering, M. and Rosmarin, K. (2000) Central beating (balanced scorecard). *Management Accounting* 78(6), 32–34.

Go, F.M. and Pine, R. (1995) *Globalisation Strategy in the Hotel Industry*. Routledge, London.

Golany, B. (1988) A note on including ordinal relations among multipliers in data envelopment analysis. *Management Science* 34(8), 1029–1033.

Golany, B. and Storbeck, J.E. (1999) A data envelopment analysis of the operational efficiency of bank branches. *Interfaces* 29(3), 14–26.

Goldberg, D.E. (1989) *Genetic Algorithm in Search, Optimisation, and Machine Learning*. Addison-Wesley, Reading, Massachusetts.

Goldstein, H. (1986) Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika* 73(1), 43–56.

Greenberg, R. and Nunamaker, R.T. (1987) A generalised multiple criteria model for control and evaluation of non-profit organisations. *Financial Accountability Management* 3(4), 331–342.

Greene, W.H. (1993) The econometric approach to efficiency analysis. In: Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds) *The Measurement of Productive Efficiency*. Oxford University Press, New York, pp. 68–119.

Grifell-Tatjé, E. and Lovell, C.A.K. (1995) A note on the Malmquist productivity index. *Economics Letters* 47(2), 169–175.

Grosskopf, S. (1986) The role of reference technology in measuring productivity efficiency. *The Economic Journal* 96, 499–513.

Grosskopf, S. (1993) Efficiency and productivity. In: Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds) *The Measurement of Productive Efficiency*. Oxford University Press, New York, pp. 160–194.

Grosskopf, S. and Valdmanis, V. (1987) Measuring hospital performance. A non-parametric approach. *Journal of Health Economics* 6, 89–107.

Grosskopf, S., Hayes, K.J., Taylor, L.L. and Weber, W.L. (1999) Anticipating the consequences of school reform: a new use of DEA. *Management Science* 45(4), 608–620.

Gstach, D. (1993a) *Comparing Structural Efficiency of Unbalanced Subsamples: a Resampling Adaptation of Data Envelopment Analysis*. Department of Economics, Working Paper No. 24, Vienna University for Economics and Business Administration, Vienna.

Gstach, D. (1993b) *Scale Efficiency: Where Data Envelopment Analysis Outperforms Stochastic Production Function Estimation*. Department of Economics, Working Paper No. 23, Vienna University for Economics and Business Administration, Vienna.

Gstach, D. (1994) *Data Envelopment Analysis in a Stochastic Setting: the Right Answer from the Wrong Model?* Department of Economics, Working Paper No. 29, Vienna University for Economics and Business Administration, Vienna.

Gstach, D. (1996) *A New Approach to Stochastic Frontier Estimation: DEA+*. Department of Economics, Working Paper No. 39, Vienna University for Economics and Business Administration, Vienna.

Haag, S., Jaska, P. and Semple, J. (1992) Assessing the relative efficiency of agricultural production units in the Blackland Prairie, Texas. *Applied Economics* 24(5), 559–565.

Hadjinicola, G.C. and Panayi, C. (1997) The overbooking problem in hotels with multiple tour operators. *International Journal of Operations and Production Management* 17(9), 874–886.

Halachmi, A. (2000) Information technology and performance measurement: promise or peril? *National Productivity Review* 19(3), 87–93.

Hanke, M. and Leopoldseder, T. (1998) Comparing the efficiency of Austrian universities. A data envelopment analysis application. *Tertiary Education and Management* 4(3), 191–197.

Harrington, H.J. and Harrington, J.S. (1995) *High Performance Benchmarking. 20 Steps to Success*. McGraw-Hill, New York.

Harris, P. (1995) Statistical cost estimation and prediction in hotels. In: Harris, P.J. (ed.) *Accounting and Finance in the International Hospitality Industry*. Butterworth-Heinemann, Oxford, pp. 83–111.

Harris, P.J. and Brown, J.B. (1998) Research and development in hospitality accounting and financial management. *International Journal of Hospitality Management* 17(2), 161–181.

Harris, P.J. and Hazzard, P.A. (1992) *Managerial Accounting in the Hospitality Industry*, 5th edn. Thornes, Cheltenham.

Harris, P.J. and Mongiello, M. (2001) Key performance indicators in European hotel properties: general managers' choices and company profiles. *International Journal of Contemporary Hospitality Management* 13(3), 120–127.

Hartley, H.O. and Rao, J.N.K. (1967) Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika* 54(1/2), 93–108.

Harville, D.A. (1977) Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* 72, 320–338.

Hawdon, D. and Hodson, M. (1996) The use of data envelopment analysis in benchmarking. *The Business Economist* 27(3), 23–40.

Holland, J.H. (1962) Outline for a logical theory of adaptive systems. *Journal of the Association for Computing Machinery* 9(3), 297–314.

Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Michigan. (2nd edn. (1992) available from MIT Press, Cambridge, Massachusetts.)

Holley, G.J. and Lynch, J.E. (1985) Marketing lessons from the UK's high-flying companies. *Journal of Marketing Management* 1(1), 65–74.

Horn, J., Nafpliotis, N. and Goldberg, D.E. (1994) A niched Pareto genetic algorithm for multiobjective optimisation. In: *Proceedings of the First IEEE Conference on Evolutionary Computation*, Vol. I. IEEE Press, New York, pp. 82–87.

Horwath International (1998) *World-wide Hotel Industry Study*. Horwath International, Coral Gables, Florida.

Horwath, E.B., Toth, L. and Lesure, J.D. (1978) *Hotel Accounting*, 4th edn. John Wiley & Sons, New York.

Hougaard, J.L. (1999) Fuzzy scores of technical efficiency. *European Journal of Operational Research* 115(3), 529–541.

Hruschka, H. (1986) Ansätze der Effizienzmessung von Betrieben. *Journal für Betriebswirtschaft* 36(2), 76–85.

Huckestein, D. and Duboff, R. (1999) Hilton Hotels: a comprehensive approach to delivering value for all stakeholders. *Cornell Hotel and Restaurant Administration Quarterly* 40(4), 28–39.

Hurley, S. and Moutinho, L. (1996) Approximate algorithms for marketing management problems. *Journal of Retailing and Consumer Services* 3(3), 145–154.

Hurley, S., Moutinho, L. and Witt, S.F. (1998) Genetic algorithms for tourism marketing. *Annals of Tourism Research* 25(2), 498–514.

Icenogle, M.L., Perdue, L.J. and Rue, L.W. (1998) Performance in private clubs: criteria utilized by managers. *FIU Hospitality Review* 16(1), 7–17.

Ijiri, Y. and Simon, H.A. (1977) *Skew Distributions and the Sizes of Business Firms*. North-Holland, Amsterdam.

Jammernegg, W., Luptáèik, M., Nakhaeizadeh, G. and Schnabl, A. (1997) Ist ein fairer Vergleich von Data Mining Algorithmen möglich? In: Nakhaeizadeh, G. (Hrsg.) *Data Mining. Theoretische Aspekte und Anwendungen*. Beiträge zur Wirtschaftsinformatik, 27. Physica Verlag, Heidelberg, pp. 225–247.

Janikow, C.Z. and Michalewicz, Z. (1991) An experimental comparison of binary and floating point representations in genetic algorithms. In: Belew, R.K. and Booker, L.B. (eds) *Proceedings of the Fourth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, California, pp. 31–36.

Johns, N., Howcroft, B. and Drake, L. (1997) The use of data envelopment analysis to monitor hotel productivity. *Progress in Tourism and Hospitality Research* 3(2), 119–127.

Joines, J.A. and Houck, C.R. (1994) On the use of non-stationary penalty functions to solve non-linear constrained optimisation problems with GA's. In: *Proceedings of the First IEEE International Conference Evolutionary Computation*. IEEE Press, New York, pp. 579–584.

Jones, P. (1988) Quality, capacity and productivity in service industries. *International Journal of Hospitality Management* 7(2), 104–112.

Kamakura, W.A. (1988) A note on 'The use of categorial variables in data envelopment analysis'. *Management Science* 34(10), 1273–1276.

Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research* 26(4), 379–390.

Kamakura, W.A., Ratchford, B.T. and Agrawal, J. (1988) Measuring market efficiency and welfare loss. *Journal of Consumer Research* 15(3), 289–302.

Kamakura, W.A., Wedel, M. and Agrawal, J. (1994) Concomitant variable latent class models for conjoint analysis. *International Journal of Research in Marketing* 11(5), 451–464.

Kantor, J. and Maital, S. (1999) Measuring efficiency by product group: integrating DEA with activity-based accounting in a large mideast bank. *Interfaces* 29(3), 27–36.

Kao, C. (1994) Efficiency improvement in data envelopment analysis. *European Journal of Operational Research* 73(3), 487–494.

Kaplan, R.S. and Norton, D.P. (1992) The balanced scorecard: measures that drive performance. *Harvard Business Review* 70(1), 71–79.

Kaplan, R.S. and Norton, D.P. (1993) Putting the balanced scorecard to work. *Harvard Business Review* 71(5), 134–147.

Kaplan, R.S. and Norton, D.P. (1996) *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business School Press, Boston, Massachusetts.

Kaplan, R.S. and Norton, D.P. (1997) Why does business need a balanced scorecard? *Journal of Cost Management* 11(3), 5–10.

Kaplan, R.S. and Norton, D.P. (2000) Having trouble with your strategy? Then map it. *Harvard Business Review* 78(5), 167–177.

Karels, G.V. and Prakash, A. (1987) Multivariate normality and forecasting of business bankruptcy. *Journal of Business Finance and Accounting* 14, 573–593.

Karlof, B. and Ostblom, S. (1994) *Benchmarking: a Signpost Excellence in Quality and Productivity*. John Wiley & Sons, Chichester, UK.

Kinnebrock, W. (1994) *Optimierung mit genetischen und selektiven Algorithmen*. Oldenbourg, Munich.

Kirk, D. and Pine, R. (1998) Research in hospitality systems and technology. *International Journal of Hospitality Management* 17(2), 203–217.

Kluge, E.A. (1996) A literature review of information technology in the hospitality curriculum. *Hospitality Research Journal* 19(4), 45–64.

Koopmans, T.C. (1951) Analysis of production as an efficient combination of activities. In: Koopmans, T.C. (ed.) *Activity Analysis of Production and Allocation*. John Wiley & Sons, New York, pp. 33–97.

Kotas, R. (1978) The ABC of PSA: an introduction to profit sensitivity analysis. *Hotel, Catering and Institutional Management Association Journal* 74, 15–19.

Kotas, R. (1982) The European hotel: methodology for analysis of financial operations and identification of appropriate business strategy. *International Journal of Hospitality Management* 1(2), 79–84.

Kotas, R. and Kreul, L.M. (1987) *Management Accounting for Hotels & Restaurants*, 9th edn. Hayden, New Jersey.

Kotas, R. and Wanhill, S. (1981) PSA – its nature, significance and applications. *International Journal of Tourism Management* 2, 176–188.

Kottke, M. (1988) Estimating impacts of tourism. *Annals of Tourism Research* 15(1), 122–133.

Koza, J.R. (1992) *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, Massachusetts.

Koza, J.R. (1997) Introduction to evolutionary computation. In: Lakhmi, C.J. (ed.) *Soft Computing Techniques in Knowledge-based Intelligent Engineering Systems. Approaches and Applications*. Physica, Heidelberg, pp. 71–111.

Kozak, M. and Rimmington, M. (1998) Benchmarking: destination attractiveness and small hospitality business performance. *International Journal of Contemporary Hospitality Management* 10(5), 184–188.

Kozak, M. and Rimmington, M. (1999) Measuring tourist destination competitiveness: conceptual considerations and empirical findings. *International Journal of Hospitality Management* 18(3), 273–284.

Kumbhakar, S.C. (1990) Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics* 46(1/2), 201–211.

Lamla, J. (1995) *Prozessbenchmarking*. Vahlen, Munich.

Larsen, E.R., van Ackere, A. and Warren, K. (1997) The growth of service and the service of growth: using system dynamics to understand service quality and capital allocation. *Decision Support Systems* 19(4), 271–287.

Lau, K., Leung, P. and Tse, K. (1999) A mathematical programming approach to clusterwise regression model and its extensions. *European Journal of Operational Research* 116(3), 640–653.

Lee, C.W.J. (1985) Stochastic properties of cross-sectional financial data. *Journal of Accounting Research* 23(1), 213–227.

Lee, K.C., Han, I. and Kwon, Y. (1996) Hybrid neural network models for bankruptcy predictions. *Decision Support Systems* 18(1), 63–72.

Leibenstein, H. (1966) Allocative efficiency vs. 'X-efficiency'. *American Economic Review* 56(3), 392–414.

Leone, R.P., Oberhelman, H.D. and Mulhern, F.J. (1993) Estimating individual cross-section coefficients from the random coefficient regression model. *Journal of the Academy of Marketing Science* 21(1), 45–51.

Lev, B.I. and Sunder, S. (1979) Methodological issues in the use of financial ratios. *Journal of Accounting and Economics* 1(3), 187–210.

Lewin, A.Y. and Lovell, C.A.K. (1990) Editor's introduction to a special issue on frontier analysis. Data envelopment analysis. *Journal of Econometrics* 46(3/5), 3–5.

Lewin, A.Y. and Morey, R.C. (1981) Measuring the relative efficiency and output potential of public sector organisations: an application of data envelopment analysis. *International Journal of Policy Analysis and Information Systems* 5, 267–285.

Lewin, A.Y. and Seiford, L.M. (eds) (1997) From efficiency calculations to a new approach for organizing and analyzing: DEA fifteen years later. *Annals of Operations Research* 73.

Lewin, A.Y., Morey, R.C. and Cook, T.J. (1982) Evaluating the administrative efficiency of courts. *Omega, International Journal of Management Science* 10(4), 401–411.

Li, X.-B. and Reeves, G.R. (1999) A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research* 115(3), 507–517.

Liberatore, M.J. and Miller, T. (1998) A framework for integrating activity-based costing and the balanced scorecard into the logistics strategy development and monitoring process. *Journal of Business Logistics* 19(2), 131–154.

Locke, E.A. and Latham, G.P. (1990) *A Theory of Goal Setting and Task Performance*. Prentice Hall, New York.

Longford, N.T. (1986) VARCL – interactive software for variance component analysis. *The Professional Statistician* 5, 28–32.

Longford, N.T. (1987) A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 74(4), 817–828.

Löthgren, M. and Tambour, M. (1999) Productivity and customer satisfaction in Swedish pharmacies: a DEA network model. *European Journal of Operational Research* 115(3), 449–458.

Loughlin, D.H. and Ranjithan, S. (1997) The neighbourhood constraint method: a genetic algorithm-based multiobjective optimisation technique. In: Bäck, T. (ed.) *Proceedings of the Seventh International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, California, pp. 666–673.

Lynch, R.L. and Cross, K.F. (1991) *Measure Up! Yardsticks for Continuous Improvement*. Basil Blackwell, Oxford.

Madu, C.N. and Kuei, C.-H. (1998) Application of data envelopment analysis in benchmarking. *International Journal of Quality Science* 3(4), 320–327.

Maital, S. and Vaninsky, A. (1999) Data envelopment analysis with a single DMU: a graphic project-gradient approach. *European Journal of Operational Research* 115(3), 518–529.

Mann, L., Samson, D. and Dow, D. (1998) A field experiment on the effects of benchmarking and goal setting on company sales performance. *Journal of Management* 24(1), 73–96.

Marcussen, C.H. (1998) Extranets of national tourist boards: Denmark, Australia, Austria, and Norway. In: Buhalis, D., Tjoa, A.M. and Jafari, J. (eds) *Information and Communication Technologies in Tourism*. Springer, Vienna, pp. 103–119.

Martin, J.C. and Román, C. (2001) An application of DEA to measure the efficiency of Spanish airports prior to privatization. *Journal of Air Transport Management* 7(3), 149–157.

Mathiyalakan, S. and Chung, C. (1996) A DEA approach for evaluating quality circles. *Benchmarking for Quality Management and Technology* 3(3), 59–70.

Mazanec, J.A. (1986a) A decision support system for optimising advertising policy of a national tourist office. Model outline and case study. *International Journal of Research in Marketing* 3(2), 63–77.

Mazanec, J.A. (1986b) Allocating an advertising budget to international travel markets. *Annals of Tourism Research* 13(4), 609–634.

Mazanec, J.A., Schmidl, J. and Wöber, K. (1996) *Zukünftige Erfolgsfaktoren im Gastgewerbe. Raumluftqualität, Nichtraucherschutz, Energie- und Ökologiebewußtsein, Managementfähigkeiten*. WIFI Wirtschaftskammer Österreich, Vienna University of Economics and Business Administration, Vienna.

McDonald, B. and Morris, M.H. (1984) The statistical validity of the ratio method in financial analysis: an empirical examination. *Journal of Business Finance and Accounting* 11(1), 89–97.

McDonald, B. and Morris, M.H. (1985) The functional specification of financial ratios: an empirical examination. *Accounting and Business Research* 15, 223–228.

McLeay, S.J. (1986) Student's t and the distribution of financial ratios. *Journal of Business Finance and Accounting* 13(2), 209–222.

McLeay, S.J. and Fieldsend, S. (1987) Sector and size effects in ratio analysis – an indirect test of ratio proportionality. *Accounting and Business Research* 17, 133–140.

McNair, C.J. and Leibfried, K.H. (1992) *Benchmarking: a Tool for Continuous Improvement*. Harper Business, New York.

Meeusen, W. and van den Broek, J. (1977) Technical efficiency and dimension of the firm: same results on the use of frontier production functions. *Empirical Economics* 2(2), 109–122.

Messier, W.F. and Hansen, J.V. (1988) Inducing rules for expert system development: an example using default and bankruptcy data. *Management Science* 34(12), 1403–1415.

Mia, L. and Patiar, A. (2001) The use of management accounting systems in hotels: an exploratory study. *International Journal of Hospitality Management* 20(2), 111–128.

Michalewicz, Z. (1996) *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd edn. Springer, Berlin.

Michalewicz, Z. and Fogel, D.B. (2000) *How to Solve It: Modern Heuristics*. Springer, Berlin.

Michalewicz, Z. and Nazhiyath, G. (1995) Genocop III: a co-evolutionary algorithm for numerical optimisation problems with non-linear constraints. In: *Proceedings of the Second IEEE International Conference Evolutionary Computation*. IEEE Press, New York, pp. 647–651.

Michalewicz, Z. and Schönauer, M. (1996) Evolutionary algorithms for constrained parameter optimisation problems. *Evolutionary Computation* 4(1), 1–32.

Min, H. and Min, H. (1996) Competitive benchmarking of Korean luxury hotels using the analytic hierarchy process and competitive gap analysis. *Journal of Services Marketing* 10(3), 58–72.

Min, H. and Min, H. (1997) Benchmarking the quality of hotel services: managerial perspectives. *International Journal of Quality and Reliability Management* 14(6), 582–597.

Mitchell, M. (1996) *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, Massachusetts.

Monarcz, E.S. and O'Brian, W. (1988) Teaching financial management with electronic spreadsheets. *Cornell Hotel and Restaurant Administration Quarterly* 29(2), 62–66.

Monarcz, E.S. and Portocarrero, N.J. (1986) *Financial Accounting for Hospitality Management*. AVI Publishing Co., Westport, Connecticut.

Monkhouse, E. (1995) The role of competitive benchmarking in small- to medium-sized enterprises. *Benchmarking for Quality Management and Technology* 2(4), 41–50.

Morey, R.C. and Dittman, D.A. (1995) Evaluating a hotel GM's performance. A case study in benchmarking. *Cornell Hotel and Restaurant Administration Quarterly* 36(5), 30–35.

Morey, R.C. and Dittman, D.A. (1997) An aid in selecting the brand, size and other strategic choices for a hotel. *Journal of Hospitality and Tourism Research* 21(1), 71–99.

Morey, M.R. and Morey, R.C. (1999) Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. *Omega, International Journal of Management Science* 27(2), 241–258.

Morey, R.C., Fine, D.J. and Loree, S.W. (1990) Comparing the allocative efficiencies of hospitals. *Omega, International Journal of Management Science* 18(1), 71–83.

Mota, S., Benzecry, J.H. and Qassim, R.Y. (1999) A model for the application of data envelopment analysis in activity-based management. *International Journal of Technology Management* 17(7/8), 861–868.

Motwani, J., Kumar, A. and Youssef, M.A. (1996) Implementing quality management in the hospitality industry: current efforts and future research directions. *Benchmarking for Quality Management and Technology* 3(4), 4–17.

Mullen, M. (1998) Rules of thumb: help or hindrance? *Accountancy* (March), 59–60.

Murthi, B.P.S., Srinivasan, K. and Kalyanaram, G. (1996) Controlling for observed and unobserved managerial skills in determining first-mover market share advantages. *Journal of Marketing Research* 33(3), 329–336.

Nanni, A.J., Dixon, J.R. and Vollmann, T.E. (1992) Integrated performance measurement: management accounting to support the new realities. *Journal of Management Accounting Research* (Autumn), 1–19.

Nissen, V. (1995) An overview of evolutionary algorithms in management applications. In: Biethahn, J. and Nissen, V. (eds) *Evolutionary Algorithms in Management Applications*. Springer, Berlin, pp. 44–97.

Norman, M. and Stoker, B. (1991) *Data Envelopment Analysis. The Assessment of Performance*. John Wiley & Sons, Chichester, UK.

Nunamaker, R.T. (1985) Using data envelopment analysis to measure the efficiency of non-profit organisations: a critical evaluation. *Managerial and Decision Economics* 6(1), 50–58. (Comment by Ahn, T.S., Charnes, A. and Cooper, W.W. in 9(3), 251–253, and reply by Nunamaker, R.T. in 9(3), 255–256.)

Nyman, J.A. and Bricker, D.L. (1989) Profit incentives and technical efficiency in the production of nursing home care. *Review of Economics and Statistics* 71(4), 586–594.

Odom, M. and Sharda, R. (1990) A neural network model for bankruptcy prediction. In: *Proceedings of the IEEE International Conference on Neural Network*, Vol. 2. IEEE Press, New York, pp.163–168.

Ogden, S.M. (1998) Benchmarking and best practice in the small hotel sector. *International Journal of Contemporary Hospitality Management* 10(5), 189–191.

Olesen, O.B. and Petersen, N.C. (1996) A presentation of GAMS for DEA. *Computers and Operations Research* 23, 323–339.

Olsen, M., Bellas, C. and Kish, L.V. (1983) Improving the prediction of restaurant failure through ratio analysis. *International Journal of Contemporary Hospitality Management* 2(4), 187–193.

Pannell Kerr Forster (1998) *A Review of the Hotel Sector in Copenhagen Compared to Selected European Competitor Cities*. Danish Tourist Board, Copenhagen.

Parkan, C. (1996) Measuring the performance of hotel operations. Socio-Economic Planning Sciences. *The International Journal of Public Sector Decision-Making* 30(4), 257–292.

Pastor, J.T., Ruiz, J.L. and Sirvent, I. (1999) A statistical test for detecting influential observations in DEA. *European Journal of Operational Research* 115(3), 542–554.

Pels, E., Nijkamp, P. and Rietveld, P. (2001) Relative efficiency of European airports. *Transport Policy* 8(3), 183–192.

Phillips, P.A. (1999) Hotel performance and competitive advantage: a contingency approach. *International Journal of Contemporary Hospitality Management* 11(7), 359–365.

Phillips, P.A. (1996a) Strategic planning and business performance in the quoted UK hotel sector: results of an exploratory study. *International Journal of Hospitality Management* 15(4), 347–362.

Phillips, P.A. (1996b) Organisational Strategy, Strategic Planning Systems Characteristics, and Business Performance in the UK Hotel Sector. PhD thesis, Cardiff Business School, Cardiff University, UK.

Phillips, P.A. and Appiah-Adu, K. (1998) Benchmarking to improve the strategic planning process in the hotel sector. *The Service Industries Journal* 18(1), 1–18.

Phillips, P.A. and Moutinho, L. (1998a) The marketing planning index: a tool for measuring strategic marketing effectiveness. *Journal of Travel and Tourism Marketing* 7(3), 41–60.

Phillips, P.A. and Moutinho, L. (1998b) *Strategic Planning Systems in Hospitality and Tourism*. CAB International, Wallingford, UK.

Phillips, P.A. and Moutinho, L. (1999) Measuring strategic planning effectiveness in hotels. *International Journal of Contemporary Hospitality Management* 11(7), 349–358.

Phillips, F., Parsons, R.G. and Donoho, A. (1990) Parallel microcomputing for data envelopment analysis. *Computational, Environmental and Urban Systems* 14(2), 167–170.

Pickworth, J.R. (1987) Minding the Ps and Qs: linking quality and productivity. *Cornell Hotel and Restaurant Administration Quarterly* 28(1), 40–47.

Plank, E.H. (1973) Determinanten für Entscheidungsprozesse im Beherbergungsbetrieb. In: Ender, W. (ed.) *Beiträge zur Fremdenverkehrsforschung*. Manz, Vienna, pp. 203–220.

Porter, M.E. (1980) *Competitive Strategy: Techniques for Analysing Industries and Competitors*. The Free Press, New York.

Porter, M.E. (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*. The Free Press, New York.

Post, T. and Spronk, J. (1999) Performance benchmarking using interactive data envelopment analysis. *European Journal of Operational Research* 115(3), 472–487.

Potter, G. and Schmidgall, R.S. (1999) Hospitality management accounting: current problems and future opportunities. *International Journal of Hospitality Management* 18(4), 387–400.

Powell, D. and Skolnick, M. (1993) Using genetic algorithms in engineering design optimisation with non-linear constraints. In: Forrest, S. (ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, California, pp. 424–430.

Ray, S.C. (1988) Data envelopment analysis, nondiscretionary inputs and efficiency: an alternative interpretation. *Socio-economic Planning Sciences* 22(4), 167–176.

Ray, S.C. (1991) Resource use efficiency in public schools. A study of Connecticut data. *Management Science* 37, 1620–1628.

Razmi, J., Zairi, M. and Jarrar, Y.F. (2000) The application of graphical techniques in evaluating benchmarking partners. *Benchmarking: an International Journal* 7(4), 304–314.

Rechenberg, I. (1993) *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, 2nd edn. Frommann-Holzboog, Stuttgart.

Reinhard, S., Lovell, C.A.K. and Thijssen, G.J. (2000) Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research* 121(2), 287–303.

Resti, A. (2000) Efficiency measurement for multi-product industries: a comparison of classic and recent techniques based on simulated data. *European Journal of Operational Research* 121(3), 559–578.

Retzlaff-Roberts, D.L. (1996) Relating discriminant analysis and data envelopment analysis to one another. *Computers and Operations Research* 23(4), 311–322.

Retzlaff-Roberts, D.L. and Morey, R.C. (1993) A goal-programming method of stochastic allocative data envelopment analysis. *European Journal of Operational Research* 71(2), 379–397.

Ritchie, J.R.B. and Crouch, G.I. (2000) The competitive destination: a sustainability perspective. *Tourism Management* 21(1), 1–7.

Ritchie, J.R.B., Crouch, G.I. and Hudson, S. (2001) Developing operational measures for the components of a destination competitiveness/sustainability model: consumer versus managerial perspectives. In: Mazanec, J.A., Crouch, G.I., Ritchie, J.R.B. and Woodside, A.G. (eds) *Consumer Psychology of Tourism, Hospitality and Leisure*. CAB International, Wallingford, UK, pp. 1–17.

Roll, Y., Golany, B. and Seroussy, D. (1989) Measuring the efficiency of maintenance units in the Israeli Air Force. *European Journal of Operational Research* 43(2), 136–142.

Rosenberg, E.A., Pieters, R. and Wedel, M. (1997) Visual attention to advertising: a segment-level analysis. *Journal of Consumer Research* 24(3), 305–314.

Rousseau, J.J. and Semple, J.H. (1993) Notes: categorial outputs in data envelopment analysis, *Management Science* 39(3), 384–386.

Ruggiero, J. (1996) On the measurement of technical efficiency in the public sector. *European Journal of Operational Research* 90(3), 553–565.

Ruggiero, J. (1998) Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research* 111(3), 461–469.

Ruggiero, J. (1999) Efficiency estimation and error decomposition in the stochastic frontier model: a Monte Carlo analysis. *European Journal of Operational Research* 115(3), 555–563.

Russo, J.A. Jr (1991) Variance analysis: evaluating hotel room sales. *Cornell Hotel and Restaurant Administration Quarterly* 31(4), 60–65.

Rusth, D.B. and Lefever, M.M. (1988) International profit planning. *Cornell Hotel and Restaurant Administration Quarterly* 29(3), 68–73.

Sakawa, M. and Yauchi, K. (1999) An interactive fuzzy satisficing method for multiobjective nonconvex programming problems through floating point genetic algorithms. *European Journal of Operational Research* 117(1), 113–124.

Santos, A.R. and Dyson, R.G. (1997) *Efficiency, Size, Benchmarks and Targets for Portuguese Bank Branches: an Application of Data Envelopment Analysis*. Warwick Business School Research Paper No. 264, University of Warwick, Warwick.

Sarrico, C.S. and Dyson, R.G. (1998) *Performance Measurement in UK Universities – the Institutional Perspective*. Warwick Business School Research Paper No. 286, University of Warwick, Warwick.

Saunders, J. and Wong, V. (1985) In search for excellence in the UK. *Journal of Marketing Management* 1(2), 119–137.

Schaffer, J.D. (1984) Some Experiments in Machine Learning Using Vector Evaluated Genetic Algorithms. TCGA file No. 00314. PhD thesis, Vanderbilt University, Nashville, Tennessee.

Schaffer, J.D. (1985) Multiple objective optimisation with vector evaluated genetic algorithms. In: Grefenstette, J.J. (ed.) *Proceedings of the First International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, California, pp. 93–100.

Scheel, H. and Scholtes, S. (1998) *Stability of DEA Efficiency Scores*. Judge Institute of Management Studies Working Paper Series No. 36, University of Cambridge, Cambridge.

Scheel, H. (2000) EMS: *Efficiency Measurement System User's Manual. Version 1.3*. Working Paper, University of Dortmund, Dortmund, Germany.

Schefczyk, M. (1993) Industrial benchmarking: a case study of performance analysis techniques. *International Journal of Production Economics* 32(1), 2–11.

Schefczyk, M. and Gerpott, T.J. (1995) Ein produktionswirtschaftlicher Benchmarking-Ansatz: Data Envelopment Analysis. *Journal für Betriebswirtschaft* 5–6, 335–346.

Schifferl, E. (1996) Optimising product development procedures with genetic algorithms: a hospitality case study. In: Beracs, J., Bauer, A. and Simon, J. (eds)

Proceedings of the 25th Conference of the European Marketing Association. Volume II. European Marketing Association, pp. 2073–2080.

Schifferl, E. (1998) Adaptive product optimisation and simultaneous customer segmentation: a hospitality product design study with genetic algorithms. In: Smith, D., Steele, N.C. and Albrecht, R.F. (eds) *Proceedings of the International Conference on Artificial Neural Nets and Genetic Algorithms*. Springer, Vienna, pp. 215–218.

Schinnar, A.P. (1990) Tradeoffs between efficiency and effectiveness in management of public services. In: Ijiri, Y. (ed.) *Creative and Innovative Approaches to the Science of Management*. Quorum, Westport, Connecticut, pp. 177–190.

Schmidt, P. (1986) Frontier production functions. *Econometric Reviews* 4(2), 289–328.

Schrage, L. (1997) *Optimisation Modeling with LINDO*, 5th edn. Duxbury Press, Pacific Grove, California.

Schwartz, G. (1978) Estimating the dimension of a model. *The Annals of Statistics* 6, 461–464.

Schwärzler, H. (1999) Strategieorientierte Unternehmensführung im Tourismus mit Hilfe der Balanced Scorecard. MSc thesis, Vienna University of Economics and Business Administration, Vienna.

Schwefel, H.-P. (1994) *Numerical Optimisation of Computer Models*, 2nd edn. John Wiley & Sons, Chichester. (Translated from (1977) *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*. Birkhäuser Verlag, Basel.)

Seiford, L.M. and Thrall, R.M. (1990) Recent developments in DEA. The mathematical programming approach to frontier analysis. *Journal of Econometrics* 46(1/2), 7–38.

Seiford, L.M. and Zhu, J. (1999) An investigation of returns to scale in data envelopment analysis. *Omega, International Journal of Management Science* 27(1), 1–11.

Sengupta, J.K. (1987) Data envelopment analysis for efficiency measurement in the stochastic case. *Computers and Operations Research* 14(2), 117–169.

Sengupta, J.K. (1988) A robust approach to the measurement of Farrell efficiency. *Applied Economics* 20(3), 273–283.

Sengupta, J.K. (1990) Transformations in stochastic DEA models. *Journal of Econometrics* 46(1/2), 109–123.

Sengupta, J.K. (1992a) Non-parametric approach to dynamic efficiency: a non-parametric application of cointegration to production frontiers. *Applied Economics* 24, 153–159.

Sengupta, J.K. (1992b) A fuzzy systems approach in data envelopment analysis. *Computers and Mathematical Applications* 24(8/9), 259–266.

Sengupta, J.K. (1992c) Measuring efficiency by a fuzzy statistical approach. *Fuzzy Sets and Systems* 46(1), 73–80.

Sengupta, J.K. (1995) *Dynamics of Data Envelopment Analysis. Theory of Systems Efficiency*. Kluwer Academic Publishers, Dordrecht.

Sengupta, J.K. (1999) A dynamic efficiency model using data envelopment analysis. *International Journal of Production Economics* 62(3), 209–218.

Sengupta, J.K. and Sfeir, R. (1986) Production frontier estimates of scale in public schools in California. *Economics of Education Review* 5(3), 297–307.

Sengupta, J.K. and Sfeir, R. (1988) Efficiency measurement by data envelopment analysis with econometric applications. *Applied Economics* 20(3), 285–293.

Serrano-Cinca, C. (1996) Self organising neural networks for financial diagnosis. *Decision Support Systems* 17 (3), 227–238.

Shafer, S.M. and Byrd, T.A. (2000) A framework for measuring the efficiency of organisational investments in information technology using data envelopment analysis. *Omega, International Journal of Management Science* 28(1), 125–141.

Shang, J. and Sueyoshi, T. (1995) A unified framework for the selection of a flexible manufacturing system. *European Journal of Operational Research* 85(2), 297–315.

Sharma, S. and Mahajan, V. (1980) Early warning indicators of business failure. *Journal of Marketing* 44(4), 80–89.

Shiroma, P.J. (1996) Efficient production planning and scheduling. An integrated approach with genetic algorithms and simulation. In: Lehner, F. (ed.) *Information Engineering and IV-Controlling*. Gabler, Wiesbaden.

Siems, T.F. and Barr, R.S. (1998) Benchmarking the productive efficiency of U.S. banks. *Economic and Financial Review* (December), 11–24.

Siguaw, J.A. and Enz, C.A. (1999a) Best practices in marketing. *Cornell Hotel and Restaurant Administration Quarterly* 40(5), 31–43.

Siguaw, J.A. and Enz, C.A. (1999b) Best practices in food and beverage management. *Cornell Hotel and Restaurant Administration Quarterly* 40(5), 50–57.

Siguaw, J.A. and Enz, C.A. (1999c) Best practices in hotel operations. *Cornell Hotel and Restaurant Administration Quarterly* 40(6), 42–53.

Simar, L. and Wilson, P.W. (1999) Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research* 115(3), 459–471.

Simeone, W.J. and Li, H.C. (1997) Credit union performance: an evaluation of Rhode Island institutions. *American Business Review* 15(1), 99–105.

Singleton-Green, B. (1993) If it matters, measure it. *Accountancy* 111(5), 52–53.

Sinuany-Stern, Z. and Friedman, L. (1998) DEA and the discriminant analysis of ratios for ranking units. *European Journal of Operational Research* 111(3), 470–478.

Smith, P. and Mayston, D. (1987) Measuring efficiency in the public sector. *Omega, International Journal of Management Science* 15(1), 181–189.

Soteriou, A. and Zenios, S.A. (1999) Operations, quality, and profitability in the provision of banking services. *Management Science* 45(9), 1221–1239.

Southern, G. (1999) A systems approach to performance measurement in hospitality. *International Journal of Contemporary Hospitality Management* 11(7), 366–376.

Spendolini, M.J. (1992) *The Benchmarking Book*. Amacom, New York.

Srinivas, N. and Deb, K. (1995) Multiobjective function optimisation using non-dominated sorting genetic algorithms. *Evolutionary Computation* 2(3), 221–248.

Sueyoshi, T. (1999) DEA-discriminant analysis in the view of goal programming. *European Journal of Operational Research* 115(3), 564–582.

Sueyoshi, T., Ohnishi, K. and Kinase, Y. (1999) A benchmark approach for baseball evaluation. *European Journal of Operational Research* 115(3), 429–448.

Sundgaard, E., Rosenberg, L. and Johns, N. (1998) A typology of hotels as individual players: the case of Bornholm, Nexoe, Denmark. *International Journal of Contemporary Hospitality Management* 10(5), 180–183.

Swamy, P.A.V.B. (1970) *Statistical Inference in Random Coefficients Models*. Springer, Berlin.

Taplin, J.H.E. and McGinley, C. (2000) A linear program to model daily car touring choices. *Annals of Tourism Research* 27(2), 451–467.

Taplin, J.H.E. and Qiu, M. (1997) Car trip attraction and route choice in Australia. *Annals of Tourism Research* 24(3), 624–637.

Tarim, Ş., Dener, H.I. and Tarim, Ş.A. (2000) Efficiency measurement in the hotel industry: output factor constrained DEA application. *Anatolia: an International Journal of Tourism and Hospitality Research* 11(2), 111–123.

Thanassoulis, E. (1993) A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments. *Journal of the Operational Research Society* 44(11), 1129–1144.

Thanassoulis, E. (1995) Assessing police forces in England and Wales using data envelopment analysis. *European Journal of Operational Research* 87(3), 641–657.

Thanassoulis, E. (1999) Data envelopment analysis and its use in banking. *Interfaces* 29(3), 1–13.

Thanassoulis, E., Boussofiane, A. and Dyson, R.G. (1996) A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega, International Journal of Management Science* 24(3), 229–244.

Thomas, R.P., Barr, R.S., Cron, W.L. and Slocum, J.W. Jr (1998) A process for evaluating retail store efficiency: a restricted DEA approach. *International Journal of Research in Marketing* 15(5), 487–503.

Thompson, R.G., Langemeier, L.N., Lee, C.T., Lee, E. and Thrall, R.M. (1990) The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics* 46(1/2), 93–108.

Thompson, R.G., Singleton, F.D. Jr, Smith, B.A. and Thrall, R.M. (1986) Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces* 16(6), 35–49.

Timmer, C. (1971) Using a probabilistic frontier production function to measure technical efficiency. *Journal of Political Economy* 79(4), 776–794.

Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985) *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons, New York.

Tomkins, C. and Green, R.H. (1988) An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability and Management* 4(2), 147–164.

Tone, K. (1989) A comparative study on AHP and DEA. *Policy and Information* 13(2), 57–63.

Tucker, F.G., Zivan, S.M. and Camp, R.C. (1987) How to measure yourself against the best. *Harvard Business Review* 65(1), 8–10.

Umbreit, W.T. and Eder, R.W. (1987) Linking hotel manager behaviour with outcome measures of effectiveness. *International Journal of Hospitality Management* 6(3), 139–147.

Valdmanis, V. (1992) Sensitivity analysis for DEA models: an empirical example using public vs. NFP hospitals. *Journal of Public Economics* 48(2), 185–205.

van der Knijff, E.C. and Oosterhaven, J. (1990) Optimising tourist policy: a linear programming approach. *Regional Studies* 24(1), 55–64.

Voss, C.A., Åhlström, P. and Blackmon, K. (1997) Benchmarking and operational performance: some empirical results. *Benchmarking for Quality Management & Technology* 4(4), 273–285.

Wanhill, S.R.C. (1980) Tackling seasonality: a technical note. *International Journal of Tourism Management*, December, 243–245.

Watson, G.H. (1992) *The Benchmarking Workbook: Adapting Best Practices for Performance Improvement*. Productivity Press, Portland, Oregon.

Watson, G.H. (1993) *Strategic Benchmarking: How to Rate Your Company's Performance Against the World's Best*. John Wiley & Sons, New York.

Wedel, M. (1997) GLIMMIX. *A Program for Simultaneous Estimation of Latent Classes and Generalised Models Within Each Latent Class. Version 1.0. User's Manual*. ProGAMMA, Groningen.

Wedel, M. (2001) GLIMMIX: software for estimating mixtures and mixtures of generalized linear models. *Journal of Classification* 18(1), 129–135.

Wedel, M. and DeSarbo, W.S. (1995) A mixture likelihood approach for generalised linear models. *Journal of Classification* 12(1), 21–55.

Wedel, M. and Kamakura, W.A. (1997) *Market Segmentation: Conceptual and Methodological Foundations*. Kluwer, Dordrecht.

Wedel, M. and Kamakura, W.A. (1999) *Market Segmentation: Conceptual and Methodological Foundations*, 2nd edn. Kluwer, Dordrecht.

Wei, Q., Zhang, J. and Zhang, X. (2000) An inverse DEA model for inputs/outputs estimate. *European Journal of Operational Research* 121(1), 151–163.

Weile, D.S., Michielssen, E. and Goldberg, D.E. (1996) Genetic algorithm design of Pareto optimal broadband microwave absorbers. *IEEE Transactions on Electromagnetic Compatibility* 38(3), 518–525.

Westermann, G. (1996) *Lokaler technischer Fortschritt und intra-industrielle Strukturen. Eine Data Envelopment Analyse*. Shaker, Aachen.

Westermann, G. (ed.) (1999) *Data Envelopment Analysis in the Service Sector*. Gabler, Wiesbaden.

Wilson, N., Chong, K.S. and Peel, M.J. (1995) Neural network simulation and the prediction of corporate outcomes: some empirical findings. *International Journal of the Economics of Business* 2(1), 31–50.

Wilson, R.L. and Sharda, R. (1994) Bankruptcy prediction using neural networks. *Decision Support Systems* 11(5), 545–557.

Wöber, K. (1993) PC-gestützte Betriebsanalyse und Konkurrenzvergleich im österreichischen Gastgewerbe. In: *Konferenzbericht des 1st Internationalen Wirtekongreß*. DEHOGA, Munich, pp. 120–125.

Wöber, K. (1994–2000) *Betriebskennzahlen des Österreichischen Gastgewerbes. Bilanzjahr 1992(–1998)*. Österreichischer Wirtschaftsverlag, Vienna. (Note: regularly published since 1994.)

Wöber, K. (1996) Tourismuswirtschaft. In: *Bericht über die Situation der kleineren und mittleren Unternehmungen der gewerblichen Wirtschaft*. BMfwA, Vienna, pp. 185–235.

Wöber, K. (1997) International city tourism flows. In: Mazanec, J. (ed.) *International City Tourism: Analysis and Strategy*. Cassell, London, pp. 39–53.

Wöber, K. (1998) Improving the efficiency of marketing information access and use by tourism organisations. *Information Technology and Tourism* 1(1), 45–58.

Wöber, K. (1999) Comparing operating ratios for small and medium hotel and restaurant businesses: a decision support system using Internet technology. In: Buhalis, D. and Schertler, W. (eds) *Information and Communication Technologies in Tourism*. Springer, Vienna, pp. 238–246.

Wöber, K. (2000) Efficiency measures in benchmarking decision support systems. A hotel industry application. In: Buhalis, D., Klein, S. and Fesenmaier, D. (eds) *Information and Communication Technologies in Tourism*. Springer, Vienna, pp. 450–459.

Worthington, A.C. (2000) Cost efficiency in Australian local government: a comparative analysis of mathematical programming and econometric approaches. *Financial Accountability and Management* 16(3), 201–224.

Yeh, Q.-J. (1996) The application of data envelopment analysis in conjunction with financial ratios for bank performance evaluation. *Journal of the Operational Research Society* 47, 980–988.

Zairi, M. (1992) The art of benchmarking: using customer feedback to establish a performance gap. *Total Quality Management* 3(2), 177–188.

Zairi, M. (1996) *Benchmarking for Best Practice: Continuous Learning Through Sustainable Innovation*. Butterworth-Heinemann, Oxford.

Zenios, C.V., Zenios, S.A., Agathocleous, K. and Soteriou, A.C. (1999) Benchmarks of the efficiency of bank branches. *Interfaces* 29(3), 47–51.

Zhou, G. and Gen, M. (1999) Genetic algorithm approach on multi-criteria minimum spanning tree problem. *European Journal of Operational Research* 114(1), 141–152.

Zopounidis, C., Doumpos, M. and Matsatsinis, N.F. (1997) On the use of knowledge-based decision support systems in financial management: a survey. *Decision Support Systems* 20(3), 259–277.

Index

Accounting data 83–84, 86–88
 balance sheet 84, 90
 gain and loss statement 85, 90
 limitations 86–88
 see also Uniform Systems of Accounts
Activity-based management 145
AHRP *see* Austrian Hotel and Restaurant Panel
Artificial intelligence 156–157
Austrian Hotel and Restaurant Panel 80–98
 DEA example 123–149
 EDEA example 181–192
 heuristic model 93–98
 mixture regression example 99–118
 reporting system 88–92
 sampling of 95, 103
 strategic groups 88
 survey design 83–86
Average ratio analysis 42, 43

Balanced scorecard 28–30, 140
 definition 28
 in hotel sector 29
 strategy map 29
Bankruptcy prediction 27, 104
Bayesian information criterion 60, 109–110
Benchmarking 1–22, 137
 best-practice 2–4, 143–144, 154, 185, 191, 195
 competitive 2–4, 156
 concepts of 7–9
 decision support systems 21–22, 94, 160, 193–199
 definition 1–2
 destination 12, 15–22
 development of 1–2
 effectiveness of 6–7
 environmental 21–22
 in hospitality sector 12–14, 21, 127–129
 partners 9–10, 115, 117, 127–129, 140, 154–155, 156–157, 160, 184–188, 191
 systems 89, 121, 191, 194
  examples 21–22, 88–98
 in tourism 11–22, 127–129, 193
 types of 2–5
Best-practice benchmarking *see* Benchmarking, best-practice

Central tendency methods 42
Change 136–137
 efficiency 136–137
 productivity 136–137
 technical 136–137

Competition comparison analysis 11
Crossover 157, 160, 173–177
 simple 174–175
 single arithmetical 175–177
 whole arithmetical 174

Darwinian principle 157
Data Envelopment Analysis (DEA) 27, 34, 61–76, 123–149
 applications 62–65
 BCC model 73–76, 124
 CCR model 65–73
 computer software tools 125
 evolutionary *see* Evolutionary Data Envelopment Analysis
 example 124–145
 interactive 148–149
 internal homogeneity 138
 limitations 144, 145–149
 longitudinal 129–137; *see also* Malmquist approach
 multimodality 152–154
 multistage 126
 orientation 124, 126, 134–135, 139, 145
 sensitivity analysis 146–148
 slack optimization 70–71, 151
 stochastic 147
 uncontrollable inputs 71–73, 137–141, 146
 window analysis 34, 124, 129–133, 147
DEA *see* Data Envelopment Analysis
DEAP 125, 135, 139, 181, 189, 191
Decision making unit 138
Destination benchmarking *see* Benchmarking, destination
Discriminant analysis 145

EDEA *see* Evolutionary Data Envelopment Analysis
Efficiency 36–41, 107, 115, 126–127, 130, 134, 143–148, 155, 157, 160, 163, 166, 183–184, 188
 allocative 37–41
 economical 37–41
 frontier 39–40, 126
 index 143–144
 measurement 155, 157, 160
  non-parametric 39, 146, 148
  parametric 39, 148
  system *see* EMS
 score 41, 107, 115, 130, 137, 146–147, 163, 166, 183–184, 188
  composite 121–122
  for efficient units 124, 142–145
  versus profit 119, 126–127
  see also Superefficiency
 technical 37–41, 134, 136, 145
EM algorithm *see* Expectation-maximization algorithm
EMS 125, 139, 181
Environmental Benchmarking Tool 21–22
Evolutionary algorithm 156, 157–158
Evolutionary Data Envelopment Analysis (EDEA) 150–192
 constraint handling 162–165
 evaluation 181–192
 initialization 165–168
 main program 161–162
 operators 168–181
Expectation-maximization algorithm 52–56, 58–59, 108, 121
 example 54–56
 in GLIMMIX 108
 limitations 58–59
 see also Mixture models
External benchmarking *see* Benchmarking

Facet participation table 130–131
Fitness 160, 162, 166–168, 170
 function 167
 scaling 166, 167, 171–173
 scores 167
Frontier 60–79, 126, 143
 minimum 147
 models 60–79, 147, 162
  deterministic *see* Data Envelopment Analysis
  stochastic 76–79, 147, 162

Genetic 156–160, 162, 165, 168–181, 189
  algorithm 156–160, 162
    constraint handling 158–159
    in tourism research 156
  classifier 158
  operators 165, 168–181, 189
  programming 158
GENOCOP 156, 158–159, 162, 168, 174, 175, 177, 180
Gini coefficient 17–18
GLIMMIX 108, 110, 111, 113

HOST *see* Hotel, Opportunity System Test
Hotel 21–22, 29, 30–31, 32, 80–92, 195
  balanced scorecard applications 29
  benchmarking *see* Benchmarking, in hospitality sector
  industry in Austria 81–82
  Opportunity System Test 21–22
  performance ratios 26, 30–31, 195
  and restaurant panel data 80–92

Input 24, 25–26, 34–36, 71–73, 103, 118, 120–121, 123, 137–139, 163
  controllable 24, 103, 118
  discretionary *see* Input, controllable
  multiple 25–26, 123
  non-discretionary *see* Input, uncontrollable
  and output transformation model 24–25, 34–36
  uncontrollable 24, 71–73, 102, 118, 120, 137–139, 163
  variables 103, 120–121
Internal benchmarking *see* Benchmarking, internal

Linear programming 151, 155, 165
Log-transformation 106, 108

Malmquist approach 124, 129, 133–137
  defined 134–135
  example 135–137
Management science 145, 156
Marketing Planning Index 13
Mixture models 49–60, 99–122
  conditional 56–58, 59, 99–122
    efficiency analysis 114–115
    example 99–122
    segment identification 59, 109–112
  expectation-maximization algorithm 52–56
  limitations 58–60, 118–122
  maximum likelihood estimation 51
  unconditional 49, 50–56
  using panel data 108–109
  see also Regression models
Moderator variable 108, 138
Multiobjective problems 121–122, 159–160
Multiperiod analysis 129, 149
Mutation 157, 160, 177–181
  boundary 178
  conditional 180–181
  non-uniform 178–180
  uniform 178

NCM *see* Neighbourhood Constraint Method
Neighbourhood Constraint Method 159–160
Neural networks 156–157
Non-Archimedean infinitesimal 71, 126
Non-profit organizations 25
NPOs *see* Non-profit organizations

Optimization 155–157
Output 25–26, 28, 102, 123
  multiple 25–26, 123
  variables 102

Panel data 80, 83–86, 100–102, 107, 123
  databases 123
  mortality 101–102
Pareto optimality 65–66, 159–160

Peer 39, 126, 127, 130–133, 137, 138, 141, 185–189, 197
appearance 130–133
groups 138, 140
members 39, 126, 127, 130–133, 137, 185–189, 197
Penalty function 158, 162, 163
Performance 4–5, 16, 18–20, 23–42, 92, 117, 120, 137, 140–142, 157
definition 30
evaluation 32–34
indicators 4–5, 16, 30, 42, 92, 137
financial 30, 92
multiple 18–20
non-financial 30–31, 84
relative importance 30
measurement 23–41, 195
with balanced scorecard 28–30
comparative 31–36
competitive 34–36
in hospitality industry 23–25, 30–31
systems 24, 31
types of 30–36
in private sector 25
in public sector 25
ratios *see* Performance, indicators
targets 32, 117, 120, 140–142; *see also* Target setting
terms 36–41
Portfolio management 145
Production function 39, 42–79, 131, 133, 134, 146
Productivity 23, 43, 92, 133, 136, 137
change 136, 137
index 133, 137
ratio 43, 92
total factor 43, 137
Profitability 126, 140

Quality management 145

Random coefficient models 33–34, 46–49
fixed effects 46
mixed effects model 47
variance component 47–49
see also Regression, models
Ratio analysis *see* Average ratio analysis
Regression 33–34, 43–60, 66–67, 107, 113, 119, 121, 145, 157
difference to DEA 66–67
log-normal 45–46; *see also* Log-transformation
models 33–34, 43–60, 66–67, 107, 113, 119, 121
see also Mixture models; Random coefficient models
Reproduction 160, 168–173
Returns to scale 124, 135, 139, 146, 162–163
constant 135, 162–163
variable 124, 135, 146
see also Data Envelopment Analysis, BCC model

Sector benchmarking *see* Benchmarking, external
Selection operator 170–172
Simplex algorithm 150–151, 181, 189
Simulated annealing algorithm 179
Small-to-medium-sized enterprises 13, 21–22, 35, 87, 149, 198
Hotel Opportunity System Test 21–22
SME *see* Small-to-medium-sized enterprises
Spider plot 18–20
Standardization 194–195
Strategic business groups 87–88, 101, 112
Superefficiency 139, 142–145, 165

Target setting 120, 127–129, 140–142
Technical change index *see* Change, technical

TourMIS 89–98, 149, 191

Uniform Systems of Accounts 86–87, 195

Variable returns to scale *see* Returns to scale, variable
Variance component models 47–49
Vector Evaluated Genetic Algorithm 159
VEGA *see* Vector Evaluated Genetic Algorithm
VRS model *see* Returns to scale, variable

Window analysis *see* Data Envelopment Analysis

X-inefficiency 38

Yield management systems 155

Z-score analysis 26–28