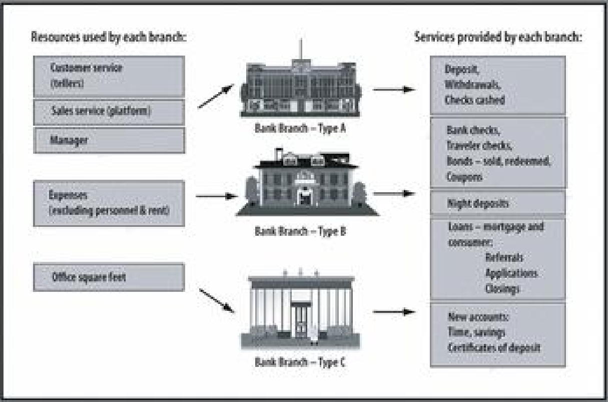
Service Productivity Management

Improving Service Performance using Data Envelopment Analysis (DEA)

H. David Sherman and Joe Zhu



DEAFrontier Software Included!

Service Productivity Management

Improving Service Performance using

Data Envelopment Analysis (DEA)

H. David Sherman
Northeastern University, U.S.A.

Joe Zhu
Worcester Polytechnic Institute, U.S.A.

Service Productivity Management

Improving Service Performance using *DATA ENVELOPMENT ANALYSIS (DEA)*

Includes *DEAFrontier* Software

Springer

H. David Sherman
Northeastern University
Boston, MA, USA

Joe Zhu
Worcester Polytechnic Institute
Worcester, MA, USA

Library of Congress Control Number: 2006922584

ISBN-10: 0-387-33211-1 (HB)
ISBN-13: 978-0387-33211-6 (HB)
ISBN-10: 0-387-33231-6 (e-book)
ISBN-13: 978-0387-33231-4 (e-book)

Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

springer.com

To Linda, Amanda and Caroline

H. David Sherman

To Yao and Alec

Joe Zhu

Contents

Foreword

By William W. Cooper

University of Texas at Austin

As the title suggests, this is a book on uses of DEA (Data Envelopment Analysis) to evaluate performances of firms in the service industries. It is more general than this, however, and better described as a user friendly introduction to DEA with examples in the service industries that can help a potential user evaluate DEA for applications that might be of interest. The applications in this book are accompanied by explanations and advice on what needs to be done to ensure success in uses of DEA.

As an introductory treatment, the book begins with a review of established methods that are already available and widely used for evaluating performance efficiencies and effectiveness. The topics covered include accounting techniques such as the use of standard costs with associated "red" or "black variances" that signal deviations below and above "efficient" performances. The discussions extend to the use of "balanced score card" approaches to determine the "effectiveness" of performances relative to goals established for programs intended to implement a corporate strategy. One shortcoming of all of these methods is that they tend to be "one at a time measures" – as is also the case for customary ratio measures such as "return on cost" or "return on net worth," etc.

By contrast, DEA simultaneously considers all inputs and all outputs that may be of interest and arrives at an overall efficiency or effectiveness score. Moreover, this is accomplished by evaluating the performance of each entity relative to a collection of entities in ways that extend commonly used "benchmark" procedures. In this way DEA identifies a subset of entities best designed to serve as benchmarks for each entity and uses them to evaluate its performance. This results in overall scores such as "90% efficiency," which means that the evaluated entity is 10% short of what it should have been able to accomplish. However, this overall score is only one aspect of what is revealed in the DEA solutions. Among other things, the sources and amounts of inefficiency in each input and output are also revealed so that a path to achievement of full (100%) efficiency is thereby obtained.

The mathematics underlying DEA models and their uses is kept to a minimum in this book. Only one of the several DEA models is formulated mathematically. The DEA literature refers to this model as the CCR (Charnes, Cooper, Rhodes) version of an "envelopment model." This name derives from the way the model "envelops" the data in order to locate a frontier where the best (i.e., 100% efficient) performers are located. This frontier is then used to evaluate the performances of other entities.

To each such envelopment model there is an associated "dual" model referred to as the "multiplier" model. This model provides further information in the form of "weights" assigned to each input and output. These weights are referred to as "multipliers" in the DEA literature in order to emphasize that they are not preassigned values like the weights customarily used in the construction of index number of prices, productivities or cost, etc. That is, the weights in DEA are determined from the data by this multiplier model for each of the entities that is evaluated.

Sherman and Zhu make extensive use of this dual (multiplier) model to increase the possibility of successful use of DEA. For example in addition to the efficiency scores, these weights can be reported for management review where it may be found that the weight assigned by the model to output A, for example, exceeds the weight assigned to output B. If this is not satisfactory it can be dealt with in a manner that does not require management to assign precise values to these weights. Instead they only need to say that they believe output B should receive a greater weight than output A. DEA can be made to take this information into account and then determine a precise numerical values for a new set of weights. The result of this recomputation can again be reviewed by management for the inefficiencies that are then identified. Also identified are new weights for all inputs and outputs. That is, in general, the changes are not confined to weights for products A and B but extend to other products as well. These results provide insights into relations between inputs and outputs that would not otherwise be apparent.

It is to be noted that DEA models provide "best estimates" for each entity being evaluated. The results therefore differ from the "average" of all performances that are used in customary statistical analysis.

Turning to more sophisticated approaches like statistical regression formulations, DEA is much less demanding in its requirements. For example, unlike ordinary statistical regressions, DEA does not require a user to stipulate the form of the relations (linear, nonlinear, etc.) that are supposed to relate inputs to outputs, etc. DEA has therefore lent itself to uses in many applications that have resisted attempts to evaluate performances by other approaches. Examples include evaluating performances of air force units or court and police performances.

In this book, these DEA properties are exploited and explained in terms of commonly used computer codes instead of the underlying mathematical models. Numerical examples and actual applications accompany the expositions. Results and managerial reactions are described that cover a variety of service industry applications.

One such example applies DEA to the numerous branches of a large U.S. bank in an application that resulted in substantial savings along with output increases and input decreases. Emphasized in this (and in other applications) is the use of quality measures, such as customer satisfactions, that play large and important roles in the service industries.

Evaluating hospitals is another example provided in this text because quality there plays a critical role. In fact, quality is a multi-dimensional and complex concept that required careful attention in the example application that is described in terms of a Health Maintenance Organization that utilizes panels of physicians ranging from general practitioners to specialists in many different specialties.

Turning to government services, an example is supplied which is based on work done with the Canadian Department of Supply and Services which, like the U.S. General Services Administration, serves as a procurer of supplies and services for the Canadian government. Here it was necessary to incorporate political considerations and to deal with initial skepticism (and even hostility in some quarters) en route to successful outcomes that are described in this book.

An emphasis on quality also appears in yet another unusual application in which DEA is used to evaluate the "quality of life" in American cities with results that compare favorably with Fortune magazine rankings (and provide added insight) based on the data obtained from the magazine's survey.

Finally, a use of DEA to evaluate "hedge fund" performances is described in order to illustrate how elements of "risk" as well as "return" can be incorporated in DEA analyses. Here the managerial use extends to "funds of hedge funds," which is to say that several hedge funds under a common fund are evaluated so that shortcomings in each of their performances may be easily identified for managerial attention.

There is, of course, much more that can be done in dealing with the service industries that now constitute nearly 80% of U.S. economic activity. This book provides a good (user friendly) start that will undoubtedly lead to further applications. Still further prospects also come into view. For instance, this last chapter (on hedge funds) provides a start toward uses of DEA in the knowledge-based-information economy toward which the U.S. is now transiting.[1]

W.W. Cooper
University of Texas at Austin
The Red McCombs School of Business
University Station 1, B6500
Austin, Texas 78712
January 2006

[1] For a good description of the transformation of the U.S. economy from manufacturing to service and now to a information based economy see Kozmetsky and Yue (2005).

Preface

How do you manage profitability of a network of hundreds or thousands of bank branches disbursed over several states and countries? How can a managed care organization manage the quality and costs of the thousands of physicians providing health services to millions of plan members? What methods would enable a government to ensure that the multiple offices serving citizens across a country are operating at low cost while meeting the required service quality? In each of these three service settings, many service providers deliver a complex set of services to a diverse set of customers. Managers of these organizations seeking excellence use benchmarking to identify and adopt best practices. Reapplying benchmarking in service organizations results in continuous improvement by allowing service units to learn from methods that have proven effective in other service units. This book presents a map of alternate ways to improve service productivity, quality, and profitability, and provides an in-depth guide to using the most powerful available benchmarking techniques to improve service organization performance--Data Envelopment Analysis (DEA).

Data Envelopment Analysis (DEA) is a proven technique to help managers improve service performance. DEA first appeared as a research tool in 1978 (Charnes, Cooper and Rhodes) and was used to evaluate educational programs for disadvantaged students in a series of large scale studies undertaken in U.S. public schools with support from the Federal government. Attention was centered on Program Follow Through—a huge attempt by the U.S. Office (now Department) of Education to apply principles from the statistical design of experiments to a set of matched schools in a nation-wide study (Charnes, Cooper and Rhodes, 1981)[1]. While DEA was originally designed for evaluating performance of non-profit organizations where market prices are not available, researchers soon realized that other organizations also had operating issues that are not adequately analyzed using financial measures. Banking seemed to be one likely business setting that could be analyzed with DEA. With bank services as the focus, DEA first appeared in management press in the *Sloan Management Review* by Sherman in 1984. Banks around the world such as Bank Boston (now Bank of America) and U.S. BankCorp, as well as brokerage firms and mortgage banks, have realized substantial cost savings from DEA since the early 1990s (O'Keefe, 1994; Iida, 1991; Bank Technology Report, 1992).

An example from the banking industry will illustrate the reason banks and other organizations have found DEA particularly valuable. Imagine that you are responsible for 1500 or more bank branches operating in 10 states in the U.S. Would you be interested in knowing that some branches are using more than twice the resources of other branches while offering the same or lower quality service than the less costly branch? DEA locates the best practice branches that are low-cost and high quality. It then identifies the high cost branches that could reduce operating costs while maintaining or improving quality. It is uniquely able to compare each one of these 1500 branches to every other branch and *simultaneously* considers all the types of resources used and all the types of transactions and services provided by each branch. Frequently the result is that there are many branches that are using substantially more resources than other branches providing a similar volume and mix of services, which was not apparent to management using other sophisticated analytic techniques. If the inefficient branches identified with DEA are aligned to approach the best practice branches identified with DEA, the bank can reduce operating costs enough to visibly increase earning per share; moreover, these savings would be annual rather than one-time benefits. Specific changes in the way services are offered, the layout of the branch and the way jobs are defined are identified through this methodology. DEA detects savings opportunities that will endure because they result from installing best practice methods derived from the best service providers in the network. This is in striking contrast to what is experienced in some organizations that achieve dysfunctional savings from across-the-board layoffs and expense cuts.

[1] For a brief background and history of DEA, please refer to chapter 1 in W.W. Cooper, L.M. Seiford and J. Zhu (2004), *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, Boston

In the above branch network, DEA identified specific excess resources in specified branches totaling over $100 million per year. Removing these resources results in real enduring savings and aligning branches with best practice branches, in contrast to cutbacks that stress the operations of the organization and which are often reversed after the crisis that spurred the across-the-board cuts subsides.

As the manager of this branch network, what would your reaction be to new information that many branches are substantially overusing resources and could operate at much lower cost without sacrificing quality and quantity of services?

- You could say you do not believe it, as you are already using some of the most sophisticated branch management techniques.
- You could try to deny it because it could raise questions about why you did not locate these excess costs earlier.
- You could conclude that it is plausible but that you will not implement these changes because your personal rewards for achieving substantial cost savings are inconsequential.
- Alternatively, you could consider the strengths and weaknesses of your current management systems and consider how this analysis complements your techniques. If it captures dimensions missing in existing analyses, which is usually the case, it will offer new opportunities to reduce costs and improve profitability.
- You also could look at examples of sets of best practice and inefficient branches identified with DEA and examine the validity of that conclusion. When managers with no knowledge of how DEA works have compared branches in this way, even where inefficient branches were reported to be using more than twice their required resources, they have found the results to be compelling, as well as, surprising.

Generally, the conclusion is that the identified weak branches *are* using excess resources, that this excess has gone unnoticed by analytic techniques in use, and that guided by DEA there are ways to improve many of the weak branches to generate substantial and enduring cost savings. All of the above responses have been encountered in applications of DEA and several of these are detailed in this book.

For several reasons, banks with large branch networks and even those with fewer than 50 branches are not able to identify the low cost branches and continue to incur waste and diminished earnings. This is the type of question that confronts Bank of America as it acquired

Fleet Boston, Royal Bank of Canada as it acquires U.S. branches, and virtually any bank with a branch network spread over many states and countries.

How can management be unaware that some branches are incurring excess operating costs compared to their own branches? First, the branches may be geographically far apart—as much as 3000 miles—and bank analytic systems (and politics) do not encourage and sometimes do not even allow such comparisons. While benchmarking with less powerful methods than DEA does occur, generally only branches in the same regions or district, or with the same manager are compared. Second, the techniques most widely used provide a sense that the branches are already well run, because 1) techniques to identify best practices cannot consider and analyze the complex set of branch transactions, and/or 2) the techniques only identify average, not the best practices.

Existing techniques such as queuing models and staffing models do not consider the complex mix of services and resources used by a bank branch and are unable to address the more complex services such as accepting loan applications and opening new accounts. Existing techniques also fail to consider the full range of services provided by the bank branch or other service unit, resulting in a splintered measure of operations and overlooking the synergies of balancing the mix of services and responsibilities assigned to service providers. These methods have been used for years and those using them have forgotten or never knew about the assumptions and weaknesses of these techniques that result in no adjustment to compensate for these weaknesses. If the individuals that evaluate and manage service costs and quality were asked what are the assumptions behind the models and methods used, it is likely that when key assumptions are identified and reconsidered, it may be surprising to realize the nature of the assumptions and how they result in sub-optimal operations. For example, widely used queuing models make assumptions about wait time and about the time it takes for each transaction. They assume each transaction is independent, which is rarely true for the majority of bank transactions.

DEA has generated substantial cost savings and productivity improvements beyond those achieved with other management techniques in diverse service organizations. Many banks have achieved costs savings with DEA while maintaining and improving service quality. DEA has identified new ways to manage health services to provide care at lower cost while maintaining quality. DEA has reduced the cost of government services in the government of

Canada. In one case, the realized benefits were so substantial that one manager described it as "benchmarking on steroids!" Users have acknowledged the benefits of DEA in news articles, journals, and professional publications. Early users had to employ internal or external experts familiar with DEA and its underlying mathematics—linear programming. The need for this type of expertise has been eliminated and the field is open for service managers seeking new insights about ways to improve their operations and competitive advantage via unique analytic capabilities of DEA. This book provides the roadmap for managers to understand DEA, evaluate its value to an organization, and apply and analyze the results to identify specific ways to improve performance and profitability.

Often new methods of managing services are identified via DEA. For example, some banks hold the view that retirement community branches and small branches are naturally less efficient—retirees like to chat with service providers and small branches have minimum staffing levels that burden their operations. DEA has been used to develop best practice small branch models that are as or more efficient than large branches. Other applications have demonstrated that there are best practice retirement community branches that can serve as models for other similar branches and they can be as productive as the large urban branches. In fact, many large urban branches have been identified with DEA as under-performing, but they were previously never challenged to improve performance because existing financial performance measures erroneously report them as best practice branches.

Similar issues confront other services environments. Both discount and retail brokerage firms have separately compared their office activities and wondered how it could identify the best practice office systems. The government of Canada has multiple offices providing similar services across Canada. Responding to the clamor for streamlining government costs of operations, they applied DEA to identify ways to provide the same service level and reduce operating costs where there is excess compared to best practice offices. In health-care, new insights about managing hospital costs, nursing home costs, and physician practice pattern costs have emerged from this technique.

This book guides managers and other readers in the use of DEA as a benchmarking technique to 1) identify high cost service units, 2) identify specific changes to each service unit to elevate their performance to the best practice services level providing high quality service at low cost, and 3) guide the improvement process. Every assertion and method described herein is supported by an application to an organization that sought to improve performance. Areas where DEA is beneficial and where it can provide only limited benefits are discussed in the context of results and managements' reaction to the DEA findings.

Most of the benefits of DEA have been discussed only in technical academic journals. Recently, this technique has become accessible to any and all managers with access to Microsoft® Excel spreadsheet software (Excel). This volume includes step-by-step guidance to enable readers to apply DEA with Excel to their organization. This book also provides ready-to-use DEA software for Microsoft® Excel Add-In to run DEA analyses on any set of organizations of interest to the reader.)

We provide an overview of the chapter contents to help readers already familiar with aspects of DEA to focus on those chapters that may offer valuable added insights.

Chapter 1 provides an overview of key existing management techniques for improving service organization performance. This suggests the landscape of methods against which a manager can evaluate the value of using DEA. Situations in which DEA would be superfluous and less insightful than other methods are identified along with examples of DEA offers superior insights but where yet other tools need to be developed.

Chapters 2 through 5 offers managers with the DEA foundation information on how to apply DEA, and how to interpret DEA results to benefit a service organization. Chapter 2 explains the basic logic behind DEA and where it is most powerful and where it is limited via simple examples of different ways to measure productivity. This chapter is ideal for those first learning about DEA. While the simple examples given may seem a little tedious and elementary, misunderstanding of basic concepts has resulted in misapplication of DEA and erroneous conclusions about ways to improve service performance. Consequently, we encourage even those managers with some background in DEA to review this chapter.

Chapter 3 offers a clear explanation of how DEA works, the computations that generate results, and the concepts that drive productivity analysis. This provides the foundation for the reader to understand appropriate ways to apply and benefit from DEA. Examples used here build in complexity, but they maintain a level where one can visualize the solution provided by DEA to allow one to anticipate its value with a full complex data-base where one could not otherwise locate the productivity improvement paths identified with DEA.

Chapter 4 introduces *DEAFrontier* software contained with this book and provides a step-by-step guidance on using this software with Microsoft® Excel to generate the solutions to the problems already analyzed in the previous chapters. This will enable a manager to run the DEA analysis to gain confidence in the software, the layout of the data and the solutions, and the interpretation of the DEA solution. This can then be adapted to any application and related data available to the reader.

Chapter 5 describes several of the many extensions to DEA. We discuss selected DEA approaches that are likely to be useful to managers, such as a returns to scale DEA model and a model that increases the power of the analysis by constraining some of elements in the analysis via input of management preferences, expertise, or other knowledge about the variables. Another particularly important DEA extension is described in depth in Chapter 7a method for including service quality in the analysis. This is an example of DEA adapting to the recognition that service quality is a critical element in measuring and managing service organizations.

Chapters 6 through 9 give detailed case studies of applications of DEA, indicating the problem being analyzed, the way DEA is applied, management's response to the analysis and the impact of the analysis. The applications include banking (Chapter 6), quality adjusted DEA (Q-DEA) applied to a bank (Chapter 7), physician practice pattern analysis (Chapter 8), and government services (Chapter 9). Three of these studies generated real documented benefits to the organization and one includes explicit management feedback on the benefits.

Chapters 10 and 11 describe new DEA uses that are in early stages of development, which have the potential to create new highly attractive options to analyze quality and to evaluate risk-reward tradeoffs in organizations. Chapter 10 describes an analysis of the quality of life in major cities using Fortune Magazine data. The concepts and value of DEA are illustrated and may form the basis for new breakthroughs in analyzing and managing service quality beyond the level suggested in Chapter 7. Chapter 11 describes assessments of hedge funds to identity those providing the greatest return for a given downside risk level. While the value of hedges fund performance measurement is substantial for investors and fund managers, the broader concept of measuring the balance of risk and reward with DEA may prove valuable to other industries that are challenged with the need to balanced risk and reward.

The authors are grateful for the comments and suggestions made by Dr. Greg N. Gregoriou on chapter 11. The authors also wish to thank Dr. W.W. Cooper for his numerous comments and suggestions on an earlier version of the book. However, any errors in the book are entirely our responsibility, and we would be grateful if anyone would bring any such errors to our attention. You are also invited to email or call with questions about the content of the book and methods of applying DEA. We would be interested in hearing about successful and problematic applications you encounter. Our objective to develop a book that would help users apply DEA and any suggestions and corrections are very welcome. Our email addresses are: h.shermand@neu.edu and jzhu@wpi.edu.

David Sherman and Joe Zhu, February 2006.

Chapter 1

MANAGEMENT OF SERVICE ORGANIZATION PRODUCTIVITY

1.1. INTRODUCTION

This chapter is designed to help service managers answer the following questions about their organization.

1. What methodologies can enhance productivity while maintaining service quality in your service organization? Based on a survey of literature and practice, we review the most widely used methods with reported benefits. Other methods may be in use that were not identified in our survey.
2. What are key characteristics of your service organization and what does that suggest about the most effective available techniques to improve and manage performance?
3. Are you using methods that are proven to be valuable in your service organization?
4. Would Data Envelopment Analysis (DEA) be a technique that would be of value in your organization?
5. How does DEA compare with other service management techniques and what are the relative strengths and limitations of these techniques?

The chapter begins with a review of basic performance terminology and concepts.

For readers interested in an explanation of basic and alternative DEA models, and computer programs, these topics are covered in Chapters 2, 3 4, and 5.

How should a service organization be managed? What skills and techniques are needed to develop a well-run service organization? How can service organization productivity be managed and improved? Interest in these questions has been propelled by growth in the service economy and greater recognition that service organizations are particularly difficult to manage with the more common and accessible management techniques. The manufacturing economy that spawned most current management literature emanated from the industrial revolution of the 19th century. That was an economy where service businesses, service organizations, and government (the largest service organization) comprised a small percent of economic activity.

Currently, over 80% of the U.S. economy is service oriented (see, Kozmetsky and Yue (2005)). Included are some of the most influential elements in our economy beginning with government, and encompassing health care, financial services, education, arts, telecommunications, law, accounting and other professional services firm, and product repair organizations. Many early management techniques that emanated from manufacturing have been adapted to service businesses such as use of standards and budgets, activity based management, production planning models, etc. Other techniques have proven effective in service organizations such as balanced scorecards, critical success factors, and six-sigma methods. The challenge of managing services is sufficiently broad and complicated to make it worth considering any new methodology that can rationalize and enhance performance. At the same time, if there were one clear comprehensive method available that ensured outstanding performance, all service organizations would adopt that method. In this book, we seek to make substantial visible headway in improving service performance, advancing the field.

1.2. PRODUCTIVITY DEFINED VIS-À-VIS EFFECTIVENESS AND EFFICIENCY

To provide a foundation for discussing methods to improve service productivity and the use of DEA, the relationship of productivity to two key performance concepts is defined, as these are universal but sometimes confused as to their boundaries. *Effectiveness* is the ability of the organization to set and achieve its goals and objectives; i.e., to do the right job. *Efficiency* is the ability to produce the outputs or services with a minimum resource level required; i.e., to do the job right. While one can address effectiveness and efficiency as separate criteria and separate foci, for a manager they are closely related and indeed efficiency can easily be viewed as a part of effectiveness. For example, a business may include as one of its goals a profit objective and achieving that objective will depend on efficiency among other things. Effectiveness will also naturally include quality objectives, which places constraints on the extent to which efficiency can be improved at the cost of quality. This quality-cost tradeoff is particularly sensitive in service organizations.

In contrast to the manufacturing environment, there are fewer objective ways of determining the quality of a service. How do you determine when an audit is really complete and up to the quality standard, whether the patient needs one more day of hospital care, or whether the advertising campaign is as successful as it should be? More resources almost always increase the cost and, if the output units are unchanged, result in lower efficiency. However, more resources can increase or decrease effectiveness. For example, frequently more hospital days and tests can be detrimental to a patient reducing the effectiveness of the care. At the same time, under treatment or under-care can damage the health of the patient as well. These are issues that drove the development of managed care in the US. (While many would agree that managed care has improved the effectiveness by providing access to the types of needed care and by limiting treatments, it is not clear that this has improved efficiency considering in the continued increase in health care costs at a rate exceeding inflation.)

Productivity is commonly defined as the ratio of outputs to inputs and is consequently focused on the efficiency of production. The terms *efficiency* and *productivity* will be used interchangeably in this book, but we note a few caveats with respect to use of these terms. First, the term efficiency can be a more "loaded" term in the sense that it may be interpreted as a value judgment of a manger's performance. While this may often be true, inefficiency is not necessarily the result of poor management performance, particularly when it is not due to elements controllable by a manager. Causes of inefficiency may be managerial, technological or socio-economic.

The term *productivity* is somewhat less sensitive, possibly because it has been less used as a value judgment term to describe a manager's performance. *Efficiency* is to some degree also viewed as having a narrower meaning i.e. you did not work hard enough or fast enough. The term "efficiency experts" has negative connotations in some environments and the term " productivity expert" has not come to be a common term though many management consultants could be characterized in that way. Finally, whichever term is used, the level of productivity/efficiency is not divorced from effectiveness issues and is specifically constrained by the type of service to be provided and the expected minimum quality that is required to allow the organization to attain its goals and objectives. Productivity management is a responsibility that can only be done well if it does not alter or violate the quality standards of the organizations. One could easily improve the productivity of any educational institution by increasing the number of students per class. This could have adverse impact on the quality of education, reputation of the institution, desirability of the institution, ability of the institution to attract donations, or other outside funding etc. This would not necessarily represent "good" management of productivity but rather "different" management resulting in higher efficiency and lower quality. It may, however, represent good management if there were adequate evidence that the reduced quality would be immaterial for all intents and purposes or if lower service quality expectations would continue to satisfy the service organization clientele or constituents.

1.3. COMPONENTS OF PRODUCTIVITY

Productivity is comprised of several components that independently influence overall organization efficiency. The components are 1) price efficiency, 2) allocative efficiency, 3) technical efficiency, and 4) scale efficiency.

Price efficiency requires purchasing the inputs that meet the quality standard at the lowest price. A service organization could increase its efficiency if it could buy its inputs (human capital and material) at a lower price without sacrificing quality.

Allocative efficiency is the use of the optimal mix of inputs to produce the products or services. This relates to capital labor tradeoffs such as the bank's use of automatic teller machines and Internet banking versus reliance on tellers or customer service representatives. This also relates to the question of whether the mix of capital equipment or the mix of labor is optimal. For example, in a hospital, there may be alternate staffing patterns which use more nursing service and fewer housekeeping services and vice versa. The optimal mix is constrained by quality standards and the relative cost of these alternative inputs.

Technical inefficiency exists when it is possible to produce more outputs with the inputs used or where it is possible to produce the actual outputs with fewer inputs. For example, one hospital found that instead of administering intravenous feeding on a decentralized basis by personnel on each floor, a centralized group of specialized personnel could provide this service. This change improved the quality of this service in terms of expertise and proved to require fewer resources resulting in a higher level of technical efficiency.

Scale efficiency is the component that addresses the optimal activity volume level. Producing more or less goods or services than the optimal level results in added costs solely due to volume or size. Inefficiencies due to scale arise due to overextended management control, operations at a level where fixed costs are high, etc. For example, a hospital will tend to need one administrator regardless of how small it is, which will tend to make the cost of administering a small hospital more costly than a somewhat larger hospital. The scale issue can, however, be easily oversimplified or misunderstood. For example, hospital research has examined scale efficiencies and has frequently focused on the number of beds as the way to measure scale size. Such analysis tends to suggest that hospitals are more economical at the 300-bed range but that scale effects are small. Regardless of whether those results are accurate for any group of hospitals, they say nothing about the optimal number of open-heart surgery cases, liver transplants, face-lifts etc. The scale efficiency issue relates to each service or type of service provided. Ignoring these facets could result in overlooking or mismanaging very influential scale efficiency dimensions.

Other components of productivity have been suggested such as financing efficiency. Rather than list other possible components, the three components discusses above, widely recognized in economic literature, essentially encompass other possible productivity components. For example, financing efficiency is largely an issue of borrowing or soliciting equity at the lowest cost. Others can convincingly argue that unless these other components are explicitly segregated, they may be overlooked.

The key point is that productivity management is a multidimensional problem and that each component requires management attention because they are independent or at least partly independent of other components. Hence, knowledge that the service organization is well-managed, efficient, or productive with respect to one component such as technical efficiency does not mean that productivity cannot be further improved by increasing price efficiency by seeking lower cost inputs.

While these may appear to be and are elementary concepts, there are numerous examples of service and other organizations that fail to explicitly recognize, analyze, and manage all of these components, resulting in missed opportunities to improve the performance of the organization. When evaluating techniques to manage productivity, it is important to explicitly consider which component of productivity are encompassed by that technique, and which components are not considered and therefore require other approaches to comprehensively manage productivity.

Multiple outputs or services provided - Many service units tend to have a diverse and changing set of services offered. Even the fast food restaurants, with contained and prescribed service procedures, are continually changing and expanding their menu to add variety and to offer different levels of nutrition. Managing a large number of services that change over time is a challenging task notwithstanding the other issues that will be discussed below. Bank tellers, for example, help process transactions, which include deposits, withdrawals, savings bond purchases and redemptions, bank check preparation etc. Each of these categories can be further segregated into other service transaction types. For example, is a deposit with one check the same as a deposit with 5, 10, or 50 checks? A Government purchasing agent may buy materials through a standing purchase order negotiated once a year and buy other materials with a one-time negotiated purchase contracts. Again, within these categories are other differences. Are purchases of armaments and defense materials different from maritime purchases, service purchases, and purchases of equipment like snow removal machines. Hospitals and physicians treat many types of patient diagnoses, which have been categorized into groups as large as over 10,000, or as small as 20-30 categories. Does each of these diagnoses need to be separately analyzed, evaluated, and managed? Alternatively are groups of these patient diagnoses similar enough to manage them as a group without unfairly or irresponsibly ignoring differences that if separately managed could lead to important productivity improvement opportunities?

Where there are many services provided and a variety of resources used to provide the services, as in the hospital setting, use of techniques that are not able to recognize and consider all the resources and services can generate erroneous views of the performance of the service providing units, such as the hospital, the physician, the brokerage branch. Data Envelopment Analysis (DEA)'s unique ability to analyze performance explicitly considering a large number of resources and services makes it particularly powerful and attractive for this type of service operation.

Multiple versus Single Location Service Provider - Managing multiple service offices or providers presents added complexity over single site services. The services offered in different locations may result in adjusting service systems to varying sets of constraints arising from customers, regulations and location differences. While most US banks had branch networks with fewer than 500 offices, US bank mergers have generated branch networks with over 2000 offices, a number that has been common in Europe and dwarfed by the State Bank of India, which manages over 8000 branches. Managed care organizations have grown to include thousands of physicians. Evaluating the productivity of these health care service providers is another challenging area that demands powerful service management techniques.

Data Envelopment Analysis (DEA) has been applied to branch networks of over 1500 offices and to managed care organizations with over 1000 physicians. It is ideally suited to compare large numbers of services, and only the computer and the software limit the number of service units that can be analyzed. While large data sets with many resources and services have been run on supercomputers, other applications have generated valuable insights with as few as seven hospitals. DEA is generally more effective with large number of units. With only a few units, benchmarking by examining the raw data and basic ratios or relationships between data elements may prove to be the most effective way to analyze the service organization performance.

Absence of efficient standards - Developing (truly) efficient standards is not practical and possibly not achievable in many service organizations. Even if efficient standards could be established for any one type of service, the numerous types of services can make this a cumbersome if not overwhelming task. This is primarily due to the professional judgments required to determine how and when a service will be provided. At one extreme, frequently encountered in health care, the professional can argue that no two services are alike due to differences in the patient and the service providers' judgments. Similar arguments are encountered if one asks a purchasing agent why two seemingly similar contracts take different amounts of time. Some banks have standards for the time required to keypunch a check document. Those same banks do not have standards for the time and resources needed to process the withdrawal, which requires the issuing of a check along with several other procedures. While one cannot generalize, several financial and government institutions that have attempted to develop service transaction standards have found that the updating of standards has been cumbersome leading to outdated and/or incomplete standards over time that limit their value as a productivity management tool. In some organizations, a poor standard may be more useful than no standard as a benchmark of productivity over time. It allows the management to track whether the service times are increasing or decreasing even if there is no basis to know what is optimal. However, there are also many cases where the tracking of standard to actual have led to complex analyses that have been too difficult to interpret and have resulted in non-use of the standard as a management tool to control productivity.

Further complicating the service management problem is the existence of activities to train, develop, and update personnel about new services, new procedures, etc. Standards for such procedures are not estimable for all intents and purposes but do nevertheless require investment of resources and therefore impact productivity.

For a standard to be meaningful, it must reflect the resources needed to produce an output. The input output relationships are not well defined for most services because of the judgment issues. If the production function, i.e. the process of converting resources into services, were known, then a standard could be established and used as a powerful tool to manage productivity. A further problem is that this difficulty in identifying the efficient production function results in attempts to use *historical* data and designating the historical mean or median resource level as the standard. These *historical standards* are not necessarily efficient standards and may result in a belief that the organization is efficient when it is really only as inefficient as it was last year. This approach prevails in areas like health care, many government activities, and many service organizations that have made serious attempts to measure and manage productivity.

Where there are objective efficient standards, DEA is unnecessary, as there are other direct ways to assess and manage the productivity of the service units. Where the standards are not objective, or based on historical actual resource levels, DEA can assist in developing standards. In short, DEA identifies best practice service providers. Studying these best practice providers can provide the basis for developing standards that reflect the most efficient methods in practice.

The bottom line is a hard and clear call of frustration that emanates from managers of all the types of service organizations. How can they manage the operation in a manner that can be defended as rational and effective? How can they be accountable for productivity with inadequate tools to accomplish this task?

Seeking ways to manage service has, out of frustration resulted in use of "seat of the pants", incomplete and arbitrary approaches that are adopted out of necessity in the absence of more comprehensive approaches. An interesting compensating result of this frustration is that we can observe these managers devoting enormous energies to find ways to manage their service organization in a way that will promote productivity and avoid many of the problems discussed above. Though it is difficult to generalize and while service organizations have a full range of management talents, these organization are not necessarily poorly managed and their productivity is not necessarily low. While more powerful management approaches can improve their productivity, competitive pressures combined with astute managers applying seasoned judgment and using techniques to motivate personnel has produced many well-run service organizations. This book endeavors to offer an approach that can assist in further improving performance for a broad set of service organizations.

1.4. TAXONOMIES OF SERVICE ORGANIZATION

There are several sets of dimensions that have been used to classify and describe service organizations. No one dimension alone proves adequate to determine the most appropriate management techniques. For example, one dimension that will be considered is professionalism. How much of the way the service is provided depends on professional or skilled judgment? One common way of identifying a professional is by their credentials, e.g., certified public accountant, attorney, engineer, and physician. This would exclude functions that do not require the same strict minimum levels of training and licensing. Nevertheless, many firms have functions that require judgments comparable in importance and complexity to those individuals in the above "Professions". Consequently, professionals also perform these other functions, where the term “professional” is more broadly defined. Examples of this would include a chef in a restaurant, a purchasing agent that is making quality versus cost tradeoff judgments in selecting a vendor, and a loan officer in a bank. The techniques appropriate to manage a professional such as a physician may be similar to those used to manage certain types of professionals and inappropriate for other types of professionals.

Ideally it would be possible to classify the service organization in a way that would help select the most appropriate management techniques. Several taxonomies for classifying service organizations have been identified from the literature and by practitioners. These can help define the more appropriate productivity management techniques. For example, recognizing that a particular government service like law enforcement cannot be evaluated with a profit measure would suggest that productivity tools that rely on profit and return on investment criteria could be inappropriate. However, other approaches like Best Practice reviews, and Zero-Base budgets may be appropriate because they do not rely on profit measures.

Classification by type of institution.

Several types of institutions can be characterized as service organizations:

A. Government
B. Non-profit organizations
C. Service businesses (service is a primary product)
D. Service businesses supporting other primary business activities
E. Internal service organizations supporting the organization and all its activities.

These are not mutually exclusive and will become apparent as each category is further described.

A. Government

Government programs, which include such diverse activities as collecting taxes, providing health and welfare benefits to particular groups of citizens and providing for national security for all citizens are comprised of numerous services. Examples of service activities found in government include

- Purchasing - procurement of a wide range of materials and services ranging from basic office supplies to defense equipment.
- Disbursement - arranging for payment for payroll, pensions for retirees, welfare benefits, vendors etc.
- Tax collection and processing

- Regulation of various activities to protect the public in areas such as securities, stock markets, drug development and testing, and transportation.
- Public Education (programs and institutions)
- Personnel programs to hire and train government employees
- Criminal justice systems and Courts of Law
- Public Health and Welfare programs and Institution such as hospitals, clinics
- Economic development programs.

The range of complexity and diversity within this partial list of government services requires a variety of management tools rather than any single tool to achieve effective and efficient operations.

B. Non-Profit Organizations

Most non-profit organizations are service oriented and share with government the objectives of maximizing program services with available resources rather than profit maximization.

Non-profit organizations include:

- Arts organizations including museums, orchestras etc.
- Educational institutions
- Hospitals, health clinics, health maintenance organizations and other health plans
- Research facilities

C. Services Businesses providing services to customers/clients for a fee as their primary business activity.

These include the professional service firms. Examples of service businesses are

- Accounting firms providing audit, tax, and consulting services
- Actuarial benefits firms
- Law firm
- Hospitals (for profit)
- Education (for profit)
- Management Consulting
- Engineering
- Architecture
- Software development
- Restaurants
- Hotels and Resorts

- Banks
- Insurance Companies
- Other Financial institution
- Airlines

D. Service Businesses that sell services directly or indirectly to customers as a secondary or support activity.

This category is segregated from services business to emphasize their existence usually within a non-service dominated business and to emphasize the need to view this as a service business even if and in spite of the fact that the primary business may be defined as manufacturing. These apparent differences between service and manufacturing may be and have often been overlooked to the detriment of the service organization that is being evaluated and managed as though it was a manufacturing business. Examples of this set of organizations include

- Customer service offices to support business products e.g. warrantee, installation, and repair service to support manufactured products such as computers, automobiles, climate control systems, etc.
- Consulting, software development, installation and training services to support the marketing and customer use of company products.
- Call centers.

E. Internal Service organization within a service or manufacturing business.

Most of these have parallel activities in the other service categories e.g. in house vs. outside legal services. They nevertheless require explicit attention in management of the business. Examples of these services include

- Purchasing
- Accounting
- Data processing
- Personnel
- In-house counsel
- Facilities management and maintenance
- Clerical support

1.5. CLASSIFICATION BY CHARACTERISTICS OF INSTITUTION

This first classification by service type immediately raises issue about which characteristics are common and peculiar to each category and what are the management implication of these similarities and differences. Several constraints, characteristics, and other cultural factors influence the options available about ways to manage productivity. One set of descriptors that capture many of these key characteristics are as follows.

1. Complexity of organization objective – constituents
2. Complexity of outputs and inputs
3. Availability and usefulness of Profit, and return on investment (ROI) measures

1. Complexity of objectives

Complexity of the organization objectives and the composition of the constituents influence the breadth of issues a manager must contend with in managing a service unit. Government and nonprofit tend to have multiple constituencies. Government constituent include the voters, legislators, receiver of services, and special interest groups. Non-profit constituents include the taxpayers that effectively subsidize these institutions, the service receiver, the donors, the board, the local and federal government, and special interest groups.

In contrast, businesses have somewhat less complex objectives in that shareholders and trustees focus on profitability and Return on Investment (ROI) subject to satisfying the specific constraints and demands of government regulations. While managing a service business is a highly complex task, the key issue is whether the shareholders are satisfied with their investment performance.. Generating continued high profitability and ROI vis-à-vis competitors is a clear comprehensive measure of successful performance for service as well as most businesses.

For example, the purchasing function in a business needs to balance price and quality as well as the tradeoff between inventory carrying cost and delivery time to procure needed material. Carrying larger inventories uses more cash but reduces the potential for losing a sale due to stock out of inventory. A lost sale is not as apparent on the financial statement as the higher inventory level that reduces return on assets and ROI.

Indeed, shareholders can challenge a business that compromises ROI to meet other objectives. For example, in 2004, the profitability of Wal-Mart, the largest retail business in the world, was criticized because they are believed to underpay their workforce and provide minimal health and other benefits. As long as Wal-Mart continues its high profitability and ROI, it need not respond to these demands. If consumers stopped purchasing goods from Wal-Mart, because of these employment practices, higher wages might be needed to retain these consumers and maintain high profitability. In contrast, some investors criticized Costco, a discount retailer that uses the club membership model, because they pay higher wages and provide more generous benefits than Wal-Mart, depriving shareholders of higher profits that could be generated with lower wages. Costco's practice may be viewed as more humane, an effective method of retaining a reliable workforce, and some may be willing to pay higher prices to support this type of management policy. Nevertheless, this practice does leave Costco at risk for a takeover by a management group that would change these practices and boost profits and ROI.

Government purchasing functions have added concerns such as support of small business, fair access of all suppliers, support for minorities, country of origin of the product vis-à-vis trade policies with foreign suppliers. Further complicating this complexity is a tilting of the priorities as the elected leaders and their appointees change.

2. Complexity of outputs and inputs

When all outputs are sold at arms-length to customers, competitive market prices for these services are generally established. These prices reflect the consumers' opinion about the value of services provided. For most service businesses, this is a key measure of the value of their activity, which allows one to compare the outputs of service organizations on a total and to some extent on a per unit basis. In comparing two banks, one that has greater fees from checking accounts might be considered superior for its greater market share, which suggests that customers prefer these services or find those fees to be more attractive. Two law firms may have different hourly billing rates, which suggest the value placed on their respective services.

In government, many services are not sold or even priced for the customer. Even when fees are established, they may not represent competitively establish market prices. What is the cost of a criminal trial, a tax audit, establishing a new law, or crime prevention? This inability to measure the output with a relatively objective quantitative measure like market price adds to the difficulty of managing service units. Similarly, internal service units within a business have similar problems. What is the value or price of computer analysis, handling a personnel matter like medical insurance inquiry, or of reviewing a contract by in-house counsel?

In the government or nonprofit setting, this is further complicated by the multiple objectives. For example, even if a certain service could be valued, is it appropriate to provide that service? While this question may be an obvious concern when one considers government health insurance to finance abortions, this is also an issue in more common question such as how much snow removal is enough and how long a wait is tolerable for taxpayers to register a motor vehicle or to obtain other licensing services.

While most inputs or resources used have market prices and are measurable for all service organizations, there are government and non-profit inputs that don't have competitive market prices. Examples of this include donated services, which may or may not have been purchased if not donated, and donated equipment. This adds to the complexity of evaluating productivity in these organizations.

Consider again the bank example. While many bank services have market values, many others of their services do not. For example, providing information on an account, visiting a safe deposit box, and cashing a check are services with no fee. Here and in the government and non-profit settings, these services need to be considered in evaluating and managing performance.

3. Accessible and meaningful measures of profit and return on Investment (ROI) measure

A business that generates what is perceived as a good return on investment over time and in comparison with its competitors would be characterized as successful with outstanding performance. This means that they are producing attractive products and services that can be sold profitably with a reasonable base of investment funds. While one could argue that a business could and should achieve greater profits and ROI though improved productivity management, these managers may counter argue that their productivity is good enough and this view may well by supported by investors, analysts, compensation consultants and bond raters. While the conclusion that high ROI and high profitability reflect outstanding performance, these measure can prevent managers from identifying inefficiencies that could further increase profits. Several DEA studies highlight the way insights from DEA provide a view behind the income statement, as will be explained in detail in later chapters. DEA identifies areas where productivity can be improved in very profitable service organizations. These productivity improvements can lead to even more impressive performance.

Where profit and ROI measures are not available, other measures of performance are often relied on. These tend to be less comprehensive and less widely accepted. The adequacy of these surrogate performance measures lends themselves to debate and managers have greater difficulty evaluating and defending their performance with these alternative performance measures. Examples of these alternatives include meeting budgeted expenses and revenues, achieving budgeted cost per unit of service, and achieving operating goals such as number of service units provided, amount of employee training provided, etc. The performance measures selected would naturally reflect key objectives of management.

The lack of profits and ROI measures primarily occurs in services organizations where market prices for outputs and/or inputs are unavailable. Government services are subject to this problem as are internal service organizations particularly where the appropriate amount of inputs cannot be objectively established. Here, the input costs are often described as discretionary costs because judgments of managers is relied on to set the amount of inputs that will be used to provide the service generated.

Where the profit and ROI measures are unavailable, managers are confronted with the frustrating task of locating alternative ways of evaluating productivity that are comprehensive, meaningful, and defensible. DEA is one of several methods that can help fill this gap in measuring and managing organization performance where profit measures are not available or meaningful.

Internal services often have no market prices making it difficult to evaluate them based on profit and ROI measures. The growth in outsourcing represents services where a market price was identified and a cost minimization decision was made which translates into profit improvement. Are data processing services, call centers, or legal services better provided in house just because they are less costly or are they better provided by outside parties that specialize in these services? When in-house service is more costly, are there other characteristics of these services that compensate or require that more costly approaches be selected over lower cost alternatives? While the answers to these questions can be highly dependent on external factors including politics and policies on international competition, the economic analyses require market prices to measure the profit impact of these options. Where these types of outsource prices are not available, management needs to seek and rely other measures of performance that may be more qualitative, judgmental, and less widely accepted.

1.6. OTHER DIMENSIONS OF THE SERVICE BUSINESS TAXONOMY

There are several other characteristics that influence the way a service organization performance can be measured and managed.

Professionalism

The role of the individual in determining the way a service is provided and resources are used to provide this service differs. This may be described as the degree of professional judgment. High degrees of professional judgment are common in the professions such as physicians, lawyers, accountants, management consultants, architects etc. Similarly high levels of judgment are associated with educational and creative activities such as artists, museum curators, and professors.

At the opposite extreme are services where unskilled labor is employed and who exercise minimal discretion and professional judgment in providing a service. Examples of this include fast food restaurants where personnel are given specific duties and where the human resources are essentially used like machinery in that the performance is relatively predictable and the amount of resources can be engineered or programmed as in a many manufacturing environments. This programming is typically accomplished through training processes and highly detailed operating procedures.

This relatively low professional judgment area lends itself well to manufacturing management techniques. This part of the service business spectrum is already well served by traditional management techniques such as standard cost systems.

The middle group of service organizations encompasses a wide range of professional judgment levels for which a pure ranking would be difficult, controversial, and also of marginal value in prescribing the most appropriate management techniques. What is important is that this middle group shares the need for management techniques tailored for the service environment. For example, dentists have more predictable methods of treatment than physicians but there remains a great deal of judgment. Loan officers also apply professional judgment within a relatively clear set of analytical approaches. Bank customer service representatives or tellers also apply judgment even though they handle a more repetitive and predictable set of transactions. While many would agree a teller has less judgment latitude than a loan officer, it would be difficult to say just how much less. It would also be difficult to agree on whether the dentist or loan officer exerts greater professional judgment.

Even at the high professional judgment level, there will naturally be wide ranges of activities. A standard will or real estate purchase and sale agreement requires less judgment by the attorney than an elaborate multibillion-dollar leveraged buyout of a public company that may include a complex range of legal issues. Hospitals care for an illness such as arthritis, cancer, or a mental disorder is highly judgmental. It is difficult to get two or more respected physician to agree on the type, amount, and timing of treatment for any one patient or any one patient type. Yet, complex life saving surgery such as coronary artery bypasses and other common heart surgery procedures have been found to be quite predictable requiring a surprisingly (at least to the laymen) narrow range of professional judgment in administering the procedure.

Management consultants that install data processing systems may also be characterized as exerting much professional judgment. Nevertheless, the consultants involved with developing corporate strategy exert yet greater levels of judgment, as there tends to be no short-term success test like "does the system work?" Hence, professionalism is a dimension that impacts the way a service firm can be managed. Yet we are not likely to have a neat rank-ordering of the amount of professionalism in various organizations.

Role of productivity

Improved productivity can benefit all organizations. There are, however, several service organizations where this is a secondary concern and others where it is immaterial for the success of the organization.

Productivity is critical to a service business like fast food restaurants that operate on low margins and where profits can be severely reduced by small increases in production costs. At the other extreme, a research organization, an experimental health care program like artificial heart implants typically use as much resources as are needed and are initially concerned primarily with the effectiveness of the program. Many government programs such as court systems also fall in this category, but there is increasing interest in explicitly considering productivity as at least a secondary level of management concern. Other examples of the kinds of organizations that place less emphasis but still consider productivity include management of artistic and educational institutions where quality of the service heavily outweighs the importance of the cost of producing the service. Professional firms paid based on time incurred, such as many law firms, are concerned with costs primarily when they appear to be excessive compared with competition. Banks naturally focus on ways to generate new funds at low cost and lending these funds at a higher rate. If a bank does not do this well, it would not survive regardless of how efficient they were in handling the deposit and loan transactions. Of late, banks have developed keen interest in managing the productivity of their operation because that also needs to be well managed to assure that they are sufficiently profitable to continue to attract investors, lenders (bond holders) etc.

Governments have also experienced the shift toward productivity consciousness. While legislators will still use as much time and resources as is needed to enact "good" laws, and while we are aware of no murder trial has yet been cancelled due to lack of funds, government programs are not automatically funded because they meet political or social needs. Those that are funded are expected to be managed in a way that minimizes waste and achieves program objectives to an acceptable degree. Today, accountability of government managers clearly extends beyond attainment of political and social program goals to the issues of productivity.

The degree to which managers will seek formalized and more powerful tools to measure and manage productivity naturally will depend on the importance of productivity in that service environment. Moreover, the lack of concern or emphasis on productivity in service organizations in the past (partly because the needed techniques were unavailable) no longer suggests that it will be relatively unimportant in the future.

Extent of customer/client contact

The degree to which services are performed independently of the customer influence the way the service organization needs to be managed. Greater autonomy and flexibility in scheduling and standardizing the processing of services occurs when customer involvement is low. Indeed, if the customer involvement can be reduced, this will give new possibilities to provide services at lower costs. Much of the move to provide customer service through the internet via a fixed elaborate set of transaction options has reduced the influence of the customer on the transaction and thereby reduced the uncertainty of resources in term of service provider time and other materials required to provide customer services. Extreme examples of the recognized impact of reducing customer involvement are cases where banks added a charge for using a human teller. The earliest application of this fee was rapidly met with a cartoon with the caption, "they want to charge me for being treated like a human being." Several airlines are also adding a surcharge for tickets purchased from an agent instead of using an Internet website. Customers are encouraged to use the options embedded in the website or incur added costs if they want to influence the available options through telephone discussion with a service representative.

Significance of quality, appropriateness, and effectiveness issues vis-à-vis productivity.

While no service organization would admit to an objective of low service quality, the relative importance of quality may be very low making productivity a dominant concern. More importantly, where quality is of key importance, productivity may be of marginal concern. For example, one test or measure of quality may be the level of customer complaints. If the customers expect low quality service or do not have the initiative to complain, then the quality objective may be easily met and the key question may be how to provide these services at lower costs. In other cases, such as health care, the quality may be managed to achieve a high minimum standard that is governed by the professionals e.g. physicians and nurses; here the question become how to provide services at lower cost while meeting this minimum quality standard.

Quality is multi-dimensional and is at least as complex as productivity management in service organizations. Until 1999, most health maintenance organizations did not measure physician quality of care. New initiatives to measure quality of hospital care, nursing home care and other health services are only gaining momentum based on initiatives that started after 2000.

The quality of audits of shareholder reports has clearly been poorly managed, which is one of the reasons one of the oldest and once highly respected firm, Arthur Andersen, evaporated in 2002. Regardless of the technical reasons for Arthur Andersen's demise, their association with many unreliable audits, including Enron and WorldCom, severely destroyed confidence in their audit opinions. The reputation of professional service firms and the quality of their service is the most important and possibly their only truly valuable asset. Impairment of their reputation is sufficient to destroy their client base. Consider whether you would return to a physician that has been found to have serious and repeating erroneous diagnoses and mistreatment of illnesses. Low quality audits can go undetected until a business encounters difficulty and the inaccuracies in reported earnings and cash flow become visible. Consequently, Sarbanes-Oxley legislation established an auditing oversight organization, the Public Company Accounting Oversight Board (PCAOB). Their initial review of the (final) big four accounting firms audits identified significant deficiencies in all of the big four's audit procedures. Moreover, there is evidence of deficiencies far beyond the level identified by the PCAOB. Pressure to improve audit quality has resulted in increasing the audit procedures, audit review processes, and expanded review of internal controls. While the fees have increased to cover added professional time required for audits, there still are no clear quality measures to assess the success of these efforts and there are questions about whether the quality improvements will prevent future debacles such as those that generated these changes (Enron and WorldCom bankruptcies).

Quality is prevalent and the dominant or primary dimension that needs to be managed in many service environments, beginning with health care and extending to government services. However, most DEA productivity studies have not addressed quality. The discussion of DEA includes a version of quality adjusted DEA (Q-DEA) which addresses this gap, provides a model to incorporate quality, and an example of the successful applications of this concept in practice.

Utilization of Equipment and other Capital assets to provide service

Professional service organizations naturally use fewer fixed asset than airlines and hospitals. Those service organizations that rely heavily on equipment naturally need to factor this resource into their productivity management approaches. Specifically, capital budgeting decisions are of key importance both to overall productivity and to human resource management in capital-intensive service organizations.

Capital budgeting techniques used in business are fully applicable to the service units both with respect to its strength and weaknesses. In addition, there are strong parallels between this type of service and robotics manufacturing in that large front end costs are incurred that will lock organization into a particular technology after which the human resources will have limited control over how the production process operates. As techniques to control robotics production evolve, they may provide useful guidance to service units that require heavy front-end fixed asset investments.

Examples in government are decisions about the type of snow removal equipment, data processing, or defense equipment to be purchased. These decisions will directly impact the number of personnel and the capabilities of personnel that will be required to provide service using these kinds of equipment. In fact, some of the gaming found in business where they lease instead of purchase equipment to preserve cash and limit reported liabilities have been found in government. The US Department of Defense proposed a lease arrangement for certain types of aircraft in 2003 which made it more palatable or politically feasible to gain approval in the budgeting process.

While selecting discount rates in corporate investment analysis requires accepting a set of judgments and estimates about the cost of debt and equity, for government projects there is added controversy about appropriate discount rate. (Generation issues appear in this controversy, where zero discount rate suggests that there is no cost and places the burden on future generations. An arbitrary 10% or other percentage may be proposed to circumvent complex arguments about the appropriate cost of capital.) Some effort to consider the timing of costs and benefits in a discounted cash flow analysis is warranted for discretionary, large dollar capital investments by governments.

Readiness, the military fourth dimension: To complete the picture of service management issues, three dimension have been considered: 1) profitability and ROI, 2) productivity or efficiency, and 3) quality or effectiveness. These three are interrelated and universal issues in service and other types of organizations. There is one other dimension that explicitly applies to defense related services that we note to complete our view of the universe of service issues. Readiness is a dimension that military and related organizations explicitly consider. For example, consider the operation of a military hospital that services aircraft carriers. This hospital may have very low occupancy during peacetime and it may provide high quality care but be highly inefficient during peacetime because it continues to maintain the ability to service an aircraft carrier full of wounded soldiers at all times. Yet, if the military is required to participate in combat, the carrier may appear with little notice with more wounded than that hospital can accommodate and it would need to be able to manage this challenge and provide the wounded with the care they need and deserve. Readiness, the fourth dimension, is one that is not well addressed and is one that is explicitly considered in the military and increasingly by businesses trying to cope with the threat of terrorism. While this term may not have been used, readiness was one dimension that was poorly managed in the government response to the significant hurricanes (such as Katrina) in 2005. Methods of developing adequate readiness and measuring this dimension are also needed for other homeland security services. It is a dimension that is not addressed in this book and one that also required development of new management techniques.

The characteristics that impact the way productivity can be managed in a service organization discussed in this chapter are summarized in Table 1-1. These characteristics will lead to a profile for any service unit that will suggest the types of productivity management techniques that are most appropriate. We now proceed to investigate the spectrum of service productivity management techniques and provide a description of their capabilities and limitations.

Table 1-1. Characteristics of Service Organizations that Influence Productivity Management

Environment complexity
Multiple objectives
Multiple constituents/customer types
Regulatory controls
Output complexity
Multiple vs. single output
Availability of a market price for each output
Rate of change in numbers and types of services provided
Multiple vs. single location service providers
Input complexity
Input choice constraints
Multiple vs. single input
Availability of input market prices
Use of equipment and other capital assets in providing service
Is an efficient standard available?
Is a profit and return on investment (ROI) measure available?
Professionalism or skill level
Importance of productivity per se in the service organization
Extent of customer contact and involvement in service delivery
Importance of quality, appropriateness and effectiveness vis-à-vis productivity
Utilization of equipment and other capital assets

1.7. SERVICE PRODUCTIVITY MANAGEMENT TECHNIQUES

Techniques widely used to manage productivity of service organizations have been identified based on a review of the literature in the fields of accounting, management control, and operations management. This set of techniques has been amplified by a nonrandom survey of approaches used by management consulting firms and techniques actually employed by service businesses. While this list is unlikely to be exhaustive, it encompasses most techniques that are used. Some techniques in use are referred to by other titles and are essentially variations on the techniques described in this chapter.

Service organizations often employ several of these techniques and different parts of the organization may be using different sets of techniques even for similar types of services. Rarely is any one approach sufficiently comprehensive and adequate for use alone. Moreover, any one or group of techniques will not necessarily be equally useful for all organizations providing similar services because of other differences in the environment, resources available, leadership style, and organization culture. Hence, this is yet another aspect of management where a contingency approach will naturally prevail. That is, the most appropriate approach must ultimately be tailored to the organization needs at the time the productivity issues are to be addressed. As is too often pointed out by consultants, if a technique is employed and is not effective, the weakness may be in selecting an inappropriate technique and the techniques may be flawed. However, another source of failures is in misapplying the technique, in which case a consultant or outside expert should be employed to improve the results next time around.

The following discussion of alternative service management techniques is offered to encourage readers to consider what type of assistance they need to improve their targeted service business.

The management tools that have been successfully used to evaluate and manage service productivity are described in this chapter. They include the following

- Standard Cost Systems
- Comparative Efficiency Analysis
- Ratio Analysis
- Profit and Return on Investment Measures
- Zero-base Budgeting

- Program Budgeting
- Best Practice Analysis or Reviews
- Data Envelopment Analysis – the focus of the balance of this book
- Peer Review
- Management Review, Management Audit, Operational Reviews, Comprehensive Audit
- Activity Analysis, Activity Based Management Functional Cost Analysis
- Process Analysis
- Staffing Models
- Balanced Scorecards (BSC)

The questions for managers to consider in reviewing these options include

- Which combination of these techniques is most appropriate?
- Are the important productivity issues addressed with this combination of techniques?
- Are there service productivity issues overlooked by these techniques where new methods need to be developed?

1.7.1 Standard Cost Systems

Standard costing is a natural place to begin. If a good standard cost system is available, most of the productivity issues can be resolved and managed without the need for other approaches.

The standard cost of a unit of service should reflect the efficient or "should be" cost of providing a service at or above a specified quality level. When this type of standard is available, the manager compares the actual cost to the standard cost for all services provided to determine whether the organization is producing these services efficiently. This analysis would also be performed for each type of service provided to determine which, if any, services are inefficient. The analysis of the variances of actual resources used from standards can provide specific information about which types of inefficiencies are present (e.g., price, mix, technical, or scale.)

This approach is dealt with at length in the management accounting literature, but the focus is skewed towards manufacturing rather than the service environment. The use of standard cost in manufacturing is fully justified by its ease of use in accounting systems to trace activity in monetary units in the production process and to assign values to inventory. Use of standards as a control tool in manufacturing might even be viewed as a secondary, albeit significant, benefit. In service organizations, the use of standards as a control tool would be the primary benefit, although standards also have applicability to service businesses for valuing unbilled services (i.e., work in process inventory). There is increasing evidence that misuse of standard cost data instead of cost analysis tailored to analyze specific business problems has led to poor or dysfunctional manufacturing decisions. Similar problems can occur and no doubt have occurred in service organizations. This is particularly important when the standard is not a true efficient or optimal standard.

The efficient or optimal standard referred to above is sometimes termed an engineered standard or engineered cost because it can be precisely determined based on a detailed understanding of the production process and the way labor, materials, and capital equipment should behave and interact to produce the specified outputs. Time and motion studies might be used to determine what resources will be required to produce a product. Other factors such as the learning curve of the organization and its employees and seasonal changes in the content of a product (e.g., cane sugar vs. corn syrup in soft drinks), will alter the efficient standard and therefore also should be considered. Management techniques to analyze manufacturing variances from standard are relatively well developed and described in most of the widely used cost accounting textbooks.

Most standard cost systems do not, however, use efficient standards and are consequently not sufficient for productivity management. A decreasing portion of manufacturing can calculate and utilize engineered standards of the type described above. Most systems use historical standards that develop the standard cost and time benchmarks based on past performance with adjustments for known changes in inputs, input costs, and production methods. The key difference is that the historical cost is not necessarily an efficient cost. The historical standard reflects actual costs in prior periods, which may include a mixture of efficient and inefficient production costs. If the efficient costs were known (i.e., which production run or set of services was efficiently generated), then the historical standard might be developed by reference to just these efficient production results. But they are often not known and are rarely known for services. Hence, past history for various time periods is usually used to establish the standard adjusted for known changes in production processes.

Comparison of actual cost to historical standard cost indicates whether operations are above or below efficiency levels of the past. Such analyses do not really indicate whether the operations are efficient, since there is no efficient standard. (A more negative view of this is that such analyses only indicate whether the level of inefficiency is above or below the historical level of inefficiency). The use of an historical standard cost system is still highly beneficial in that it highlights areas where costs are above or below the expected, budgeted, or standard level. This allows management to determine what types of actions are needed to improve control of operating costs and/or to revise their expected costs and assess the implications for the organization (e.g., change in profits, cash flow, etc.)

Widespread use of historical costs may be justified on several grounds. First, if engineered costs cannot realistically be developed, this technique represents a practical second best approach. Second, the historical cost may be considered close enough to efficient costs, although research may be needed to support this view. The potential benefit from improving efficiency over the historical standard may be so small that the historical standard is deemed to be adequate. Of course, management does not have any gauge to measure these potential benefits. Management may have yet other justifications for relying on the historical standard. This may reflect management's high regard for the middle manager's abilities to economically produce their product or service outputs and/or the reliance on competition to motivate management to operate as efficiently as their capabilities allow.

While some question the degree to which standards are used today, one of the authors completing a DEA study with one of the five largest US banks in the late 1990s observed continued development and use of standards. The management consulting firm hired by the bank was sending staff to branches with stopwatches to estimate processing times. This effort was done to develop time standards for transactions. Many banks also have such standards, which are often outdated due the cost and effort of revising these as services and procedures change. These standards are used for pricing of internal services and staffing. While there is evidence of continued use of standards, the efficacy of these standards to manage productivity is not always apparent and can be the basis for false confidence in the organization's productivity. Periodic review of standards and their benefits in managing services is warranted for organizations using this methodology.

1.7.2 Comparative Efficiency Analysis

If engineered or efficient standards are not available, the approaches used to evaluate productivity generally utilize some type of comparative efficiency analysis (CEA). Comparative efficiency analysis requires that performance be compared against judgments, opinions, past history, other organizations, etc. to assess the efficiency of operations. This situation is pervasive in service organizations where the efficient standard rarely exists. A few caveats are in order to recognize what happens in the realm of CEA.

1. The standard benchmark will have some inherent flaws. Using this benchmark to assess performance may suggest operational problems when the real problem is the flawed benchmark.
2. While the first caveat is an elementary concept -- the standard may need to be revised when variances are observed -- it can be easily forgotten or overlooked during the analytical phase. That is, the historical standard may come to be viewed as an accurate, efficient standard over time and its inherent flaws may fade in memory.
3. Sometimes historical standards are developed for purposes that are unrelated to productivity management and adapting these for productivity assessment can be misleading and even dysfunctional.

   A common example of this is that costs of services are often developed to establish transfer prices for these services. This process usually has the objective of allocating and/or billing all costs to users. It is generally understood that almost all transfer prices have somewhat arbitrary and judgmental elements. These transfer prices and their underlying costs may eventually be viewed as the standard cost, even though there was no attempt to estimate efficient cost when the transfer price was established.

With this understanding of the potential limitations of comparative efficiency analysis, we proceed to consider individual CEA techniques widely used in service and other organizations.

1.7.3 Ratio Analysis

Productivity measurement naturally evokes the concept of a ratio of outputs to inputs. If an efficient standard were available, the ratio of standard to actual resources used would represent an efficiency ratio. A ratio of actual output to efficient output equal to 1 or 100% would mean the process is operated at maximum efficiency. Where standards are not available, ratios are often used as a gauge of operating performance. Cost per transaction, cost per unit of output, amounts of resources per unit of output, units of resource A used per unit of resource B used; these are examples of ratios that may be calculated and analyzed. Often many different ratios are calculated to focus on different aspects of operations. Ratios are generally used to compare various dimensions of performance among comparable units and within a single unit over time periods.

For example, two hospitals or the same hospital over two time periods might compare cost per patient or cost per day of care. Assume Hospital A had costs of $300 per day and Hospital B has costs of $350 per day. What does this ratio tell the manager about productivity? Several possibilities exist, some of which are listed below.

- "A" is better managed and is more efficient at providing care.
- "A" treats less severe and/or less resource consuming patient illnesses.
- "A" can buy inputs at lower cost but is no more technically efficient than "B".
- "A" is operating at a volume level where there are scale economies.
- "A" does not provide program services that "B" offers such as health improvement classes, nurse training, etc.
- Some combination of the above and other reasons.

In fact, "B" with its higher costs may be more efficiently operated than "A" but has an output mix of severe cases and other services that results in a higher cost per day. Even after we have considered these factors and reach a conclusion about whether "A" is more or less efficient than "B", there is the possibility that neither is very efficient.

Beyond this is the question of whether the knowledge that "A" is more or less efficient than "B" is helpful to the manager in improving "B's" efficiency. If "B's" costs should be cut $25 per day to make it efficient, how does "B's" management determine where these reductions should be made?

To compensate for the inability of one ratio to capture the output mix differences and to segregate the types of inefficiencies, other ratios are often developed resulting in a set of ratios providing a profile of the operating results. In hospitals, the other ratios might include nurses per patient day, ratio of high-risk patients to all patients, cost per meal, housekeeping cost per bed-day, etc. This ratio profile provides insight into the components that may require attention to improve productivity. For example, if the purchase price of certain resources used is relatively high, management may focus on this potential price inefficiency for cost reductions.

Several limitations of ratios are apparent when a set of ratios is used. First, it is difficult to interpret the complexity of a set of ratios intended to evaluate the organization. For example, if the cost per patient is high and the ratio of high to low resource intensive patients is also high, the latter will explain some of the higher cost. What is not known is whether this explains *all* of the higher cost. Finally, when comparing similar organizations or one organization over several time periods, the ratios present a range of actual operating results. There is no objective point above or below which the unit is efficient or inefficient. Consequently, use of ratios to identify inefficient units requires a judgmental separation of the efficient and inefficient level. Typically, some arbitrary rule of thumb is adopted to establish this cutoff; one might say that units with costs more than one standard deviation above the mean are inefficient. Since this is arbitrary, it is difficult to defend this as a meaningful cutoff and there is no assurance that some or many of the units with costs below this cutoff are not also inefficient. This type of ratio analysis is similar to, but less precise than, the variance analysis applied to standard cost systems.

Despite some limitations, ratios are nevertheless very helpful in many instances and their use in tandem with other techniques can result in very powerful actionable insights that management can use to improve productivity. Indeed, when organizations provide only one type of service that meets the quality standard and use one type of input, ratios of output to input provide significant insights into which units are inefficient. The unit with the highest output to input ratio would be the most efficient unit (assuming there are no unusual aberrations or data problems) and units using more resources could reasonably be challenged to either meet the efficiency level of the one most efficient unit or justify their higher resource utilization.

1.7.4 Profit and Return on Investment Measures

Profitability (defined as income divided by revenue) and return on investment (ROI; defined as net income divided by invested capital) are two ratios that are widely used in analyzing business performance. They apply to services as well as manufacturing businesses. In many service businesses most of the investment is in human resources (rather than in capital equipment), which is not capitalized as an asset for financial accounting purposes but rather is reflected as a period expense. Training and hiring costs reduce current income resulting in downward pressure on ROI. The result is that ROI in many services behaves differently from manufacturing and more capital intensive businesses. . This difference is not important to investors who focus on the risk and return on monies they invest in the business and expect a fair ROI on all of their investments. This difference is, however, of importance to managers who may be responsible for both service and manufacturing activities, since these differences need to be understood to evaluate, manage, and allocate resources among these business units.

The profitability and ROI ratios both have limitations and benefits in business performance evaluation, as discussed below.

1. The ROI measure is a comprehensive measure in contrast to profitability, which ignores the amount of invested funds used to generate profits. High ROI <u>over</u> <u>time</u> reflects effective <u>and</u> efficient operations.
2. Comprehensive measures like ROI are more meaningful than the profit measure for overall performance evaluation. Consequently, ROI is widely used to measure business performance that is the type of measure a manager might seek for government and non-profit organizations, if possible.
3. Both profitability and ROI are potentially subject to a short-run bias (i.e., this year's performance can appear to be strong by sacrificing long term performance). For example, delayed training and hiring or deferred development of new service

products will boost current profits but delay benefits from these activities. Even the more comprehensive measures like ROI need to be constrained to balance short-term and long-term performance. Resource consuming activities like training and marketing that benefit future periods will tend to make ROI appear lower but this does not necessarily represent reduced productivity.

4. Increased productivity will tend to increase profitability and ROI over time. <u>However, ROI and profitability ratios will not necessarily be sufficient to locate areas of poor or outstanding productivity</u>. For example, banks that are very profitable may be unaware of significant cost-saving possibilities, which could further boost profits.

1.7.5 Zero-base Budgeting

Zero-base budgeting (ZBB) and the subsequent topic, program budgeting, are strongly associated with government. Many of the concepts apply to other organizations that produce services, and are particularly valuable where there are no comprehensive measures of profitability and where there are not objective market prices for services provided. Departments within an organization, such as human resources and information systems, share the characteristic of not selling outputs and not having objective market prices may enhance their performance through use of ZBB and program budgeting.

Service organizations that lack the standards needed to develop a budget that objectively determines the revenues and expenses nevertheless develop budgets as a planning tool to reflect their operating plans to estimate financial resources needed and as a control tool to measure whether the organization is achieving its goals and where adjustment are needed to attain these goals. Because the amount of funds used is largely based on management judgment, they are often referred to as discretionary costs. The actual results are compared with the budget to understand where and why they differ from the budgeted figures and to assess the extent to which the organization is meeting its operating plans. The variances of actual results from the budget can be used by management to activities that require added management guidance, areas of where the performance exceeds expectations which the organization might expand and other changes needed to achieve organization objective. The budget is typically based on historical cost levels adjusted for known changes in the cost structure or procedures. Achieving this budget does not necessarily assure a high level of productivity.

Zero-base budgeting (ZBB) or a zero-base review is an approach to help managers develop budgets where standards are not available. The objective is to ensure that the budgeted cost level is rational and efficient. In essence, ZBB is a mechanism by which management can closely examine and reconsider the resource levels used by operating units for which there are minimal benchmarks regarding what is efficient and effective performance.

Managers of each department under ZBB are asked to separate their department activities into decision units. Each decision unit assembles a decision package, which outlines the functions, goals, and costs of the decision unit. Managers must devise various programs, which will enable the decision unit to attain its goals, and these programs are included in the decision package with the most feasible program identified as such and the alternatives listed.

Each decision package is evaluated in light of the costs and benefits of the proposed program and then ranked in order to select those that are most important to the organization. The ranking process will determine where resources will be allocated. Resources are typically allocated to the highest ranking decision packages. The benefit of this approach is that it focuses management attention on a review of what functions are performed and why, as well as the costs and benefits of these functions to the organization. After considering the value of each of the ZBB budget packages with respect to the current objectives of the organization and resource constraints, management must approve or reject each package. The result of this review may be to increase or decrease the activity and cost level. Hence, ZBB does not necessarily lead to increased productivity or reduced costs. Rather, ZBB seeks to balance productivity, quality and other management priorities to achieve the organization's goals. While ZBB analysis often is judgmental and qualitative , it is designed to help management match their priorities with the resources allocations and thereby rationalize the budget. The process of ZBB forces management to justify the expenditures under review for the upcoming year rather than simply applying an inflation adjustment to the prior year budget. ZBB has been subject to criticism because of the time and other costs associated with it and due to some questions about its ultimate effectiveness. Nevertheless, several large corporations have found ZBB to be beneficial in managing internal staff functions and have described it as a preferred tool.

Variations on ZBB have been adopted to pressure management to improve productivity. For example, management could increase the annual department budget by inflation less some fixed percent (i.e., inflation minus 2%). Improved productivity may result particularly if there have been no efforts to manage productivity in the past. Nevertheless, this approach can also result in a cutback in service with no change in productivity, which may be a dysfunctional result. In cases where variable costs are 50% of the total cost, a 1% cost reduction in variable costs may result in a 2% reduction in service resulting in a more dramatic result than was anticipated. It also may be demoralizing to departments that are already highly efficient and cannot readily reduce their budget level without incurring negative side effects, such as decreased quality of service. ZBB represents one way to avoid such across-the-organization arbitrary pressures to reduce cost.

Application of ZBB has evolved into a focus on the expenditures above a certain point, such as 50% of the prior budget. This allows required and statutory activities over which management has little control to be automatically approved without specific review. In addition, the substantial effort required for a detailed ZBB review has led to its use on a cyclical basis where each service is reviewed every 3 to 5 years (see, for example, Pyrrh (1973) and Brown (1981) for further discussion of ZBB). In addition, to manage the effort and costs of ZBB, it can be cycled to different departments of the organization so that only a small percentage of the organization is subject to this analysis in any one year.

An example of a variation of ZBB used by a local non-profit radio and TV station illustrates the way the concept might be adapted to a service organization's objectives. They request that each departmental report on what would change if the department budget was cut 5% and 10%, and what would be added if the budget were raised 5%. This process is used to determine which sets of budgetary changes are most consistent with the organizations current objectives. Similar adaptations have been used by museums and other non-profit organizations.

ZBB is most appropriate for service areas where little or no direct revenue is generated and when it is most difficult for managers to determine the efficient and effective amount of resources needed to meet the service volume and quality objectives. Areas where revenues and/or more specific output units can be measured are more effectively managed through program budgeting techniques as described below.

1.7.6 Program Budgeting

Sometimes referred to as Program Planning and Budgeting Systems (PPBS), program budgeting is designed to indicate the resources used to provide specific services or groups of services identified as programs of the organization. A comparison of the program costs with its estimated benefits is the basis for reallocating, augmenting, diminishing, adding or deleting programs. Management defines the programs based on their goals and objectives and analyzes them consistent with these definitions. For example, a graduate business school might treat the masters of business (MBA) as the program. Another business school that has a lock step first year and electives in the second year might define the first year of the MBA as a program. Similarly a hospital might define organizational department like radiology and operating rooms as programs. Alternatively, they may view all elder care as a program resulting in combining the costs and revenues from several departments to describe the economic value of the program.

The objective of program budgeting is to assess each program by evaluating the resources used in comparison with the revenue generated or the services provided. Program budgeting indirectly affects the productivity of an organization in several ways. The primary is improved resource allocation to augment the benefits generated with available resources. Program analysis may suggest areas where the cost of the program is not justified by the benefits. Such programs may be discontinued or scaled down. Resources may be redeployed from discontinued programs to increase the productivity of other programs. Additionally, program budgeting can improve productivity by segregating one program's costs and benefits from those of other programs and activities in the organization, which can sharpen the focus on ways to improve the management of each program. While program budgeting is not solely focused on productivity, it generates actions to enhance productivity. In a service business that sells its services to outside customers, a program budget analysis would be a product line analysis for those marketed services.

For example, an accounting firm might consider the profitability and long-term benefits of its audit product line (or program) versus its tax services. Within the audit area, it may further consider specific industry audit profitability. A program review of non-commercial services in the same firm might be focused on training programs, practice development marketing programs, recruiting programs, etc.

Program budgets lend themselves to direct correspondence with responsibility assignments in an organization. Consequently, this can be an effective tool to monitor and control operations over time. This requires that the accounting system be adapted to the program structure so that revenues and costs can be captured and reported on a program-by-program basis. Adding the program codes to each transaction and revising coding procedures accordingly can readily accomplish adapting accounting systems to a program budget. In this way, the program budget becomes a standard against which changing productivity can be measured and analyzed to some extent.

1.7.7 Best Practice Analysis

When the efficient standard is unavailable and an historical standard is either not feasible or unlikely to be reliable, a "best practice" review approach may prove to be a useful alternative. This is a process by which individuals, groups, or organizational units that provide similar services compare their operating methods, outputs, and resource utilization to establish a standard that will be used as a benchmark for efficient operations. The best practice standard represents the collective views of service providers about the most effective and efficient way to provide this service. It is the result of analysis, discussion, and to some extent negotiations about how the service should be produced. Essentially, this is an attempt to develop a standard that is efficient in resource utilization and effective in meeting quality and service objectives of the organization. It requires the sharing of information among service providers, and consequently it is less likely to be implemented across competing business organizations. This approach does lend itself to governments where the nature of competition does not tend to limit data sharing and where much of the information is already in the public domain.

Professionals such as health care providers have employed best practice analysis to create treatment protocols for specific illnesses. These have been formally adopted hospitals as their “standard of care". Here the objective is to ensure effective care while containing costs. If a physician or hospital staff follows this standard care, the patient is believed to be receiving reasonable treatment. If a patient experiences an adverse reaction that triggering a malpractice law suit, adherence to the standard of care provides some evidence that the treatment is within norms condoned by health professionals. The standard of care also provides one basis for negotiating fees for health services. Of course fee negotiations and malpractice claims are also influenced by many other factors. Even these standards are not necessarily reflecting the most efficient or effective way to treat a patient, which is why hospital standards of care will differ and why physicians may choose a treatment pattern that differs from the standard if they believe their patient requires another treatment pattern.

Best practice reviews are costly in terms of human resource time required. However, the result is a standard that each participant's service units will strive to achieve. Inasmuch as it is based on actual operating experience, this standard will be achievable, although significant operating changes in each organization may be required.

The most serious limitation is that the benchmark developed is a best practice rather than an efficient one. It is not explicitly designed to locate new ways to provide services with new technologies. Such possibilities may, however, be identified via this process and these findings can yield significant productivity improvement.

The key benefit of best practice reviews is their ability to address service activities for which it is extremely difficult to develop standards because of heavy reliance on judgment and situation specific characteristics in determining how a service will be provided. One example of this approach is the development of medical care cost and procedure standards by physicians within groups of U.S. hospitals. This has provided significant insight about what various illnesses should cost for treatment.

This best practice review process also occurs informally and formally within other service organizations such as architectural firms, money management firms, and software development firms. Meetings held to critique and develop ways to achieve effective results at minimum cost are frequent occurrences. Formalization of this process with a focus on productivity can result in a very specific best practice profile as has been achieved in the health care setting in selective applications to date.

1.7.8 Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA), a quantitative technique is used to establish a best practice group of units and to determine which units are inefficient compared to the best practice groups and the magnitude of inefficiencies present. Consequently, DEA clearly and objectively indicates which units should be able to improve productivity and the amount of resource savings and/or output measures that these inefficient units must achieve to meet the level of efficiency of the best practice units.

Through use of linear programming, DEA obtains these insights by comparing the set of actual outputs achieved with the set of actual inputs used by organization units providing a similar set of services but naturally with varying volume and mix of services. It directly incorporates multiple inputs and outputs, which means that the results will be explicitly sensitive to the complexity and mix of outputs. It can incorporate outputs that have no clear price or market value, like training and new service development, which will benefit future periods but which consume resources in the short-term. Like the best practice review, it compares units that provide similar services; it does not locate new technologies to improve productivity except where individual units have adopted new technologies that make them relatively efficient. Unlike the best practice review, it is highly objective and focuses primarily on technical and scale efficiency. A key attribute is fairness in that the units that are found to be inefficient are located after considering their mix of inputs and outputs. A second benefit is that it provides strong indications of what type and amount of changes in inputs and outputs are needed to make inefficient units efficient.

The basics of DEA will be described in detail in Chapter 2. At this point, the primary characteristics that are of note are that it is a best practice technique that can objectively locate real productivity improvement possibilities without the need for any standards and that it identifies best practice and inefficient units by comparing their actual operating results. DEA does require data on resources used, services provided, and other descriptive data be available in contrast to best practice reviews and peer reviews that can be completed with qualitative data. (We should note that qualitative data or information could be incorporated into DEA. See, e.g., Cook (2005) and Cook and Zhu (2006).)

1.7.9 Peer Review

Peer reviews are designed to allow an organization to benefit from the knowledge and expertise of outside professionals who are experienced in providing specific services. These consultants have a broad scope of knowledge in their fields, which enables them to address quality, effectiveness, and efficiency dimensions. In the absence of an objective benchmark of efficient and effective management, peer reviews can help management examine how an organization can improve the way in which it provides and manages its services. The input of qualified professionals provides information that may improve the efficiency of the organization. Increased productivity is only one of several possible outcomes. For example, in some cases physician peer reviews have resulted in decreased length of hospital stays leading to increased productivity (Churchill, Cooper and Govindarajan (1982)). In other cases additional patient tests, added treatments, and extended hospital stays have been suggested which improve the quality of service and reduce the risk of malpractice but at the same time decrease productivity by increasing costs. While productivity will not necessarily improve or approach the maximum attainable level as a result of this review, it does provide management the assurance that the issues of productivity and quality have been explicitly considered and that the organization has had the opportunity to benefit from the perspective of highly qualified professionals. To the extent that these professionals are recruited from without the organization, the peer review may be relatively objective and representative of the best professional judgments drawn from a cross-section of organizations.

1.7.10 Management Reviews, Management Audits, Operational Reviews, and Comprehensive Audits

This set of processes is tantamount to an elaborate peer review by independent individuals outside the organization unit. This differs from, and is more comprehensive than, a peer review in many cases because the types of expertise employed may go well beyond the service provider's profession per se. The team of individuals conducting such a review will include not only qualified peers but also individuals qualified in other analytical techniques needed for the review, such as survey experts to collect customer satisfaction data, statistical experts to analyze the data collected in the review, and human resource management experts to evaluate the way professionals are managed and motivated. Management audits also differ from peer reviews in that results of the audit are presented to a third party, usually a more senior management group, in addition to the unit being audited. This design results in greater accountability for responding to the findings than a more insulated peer review. During a management review the independent auditors or reviewers analyze financial, operational, and managerial performance to assess the efficiency and effectiveness of the organization. The analysis may include observing operations, reviewing both internal and external financial statements and other operating data, and discussing areas of concern with personnel. The credibility and objectivity of the management audit are often greater than with the peer review because of the independence of the auditors and the broad range of expertise brought to bear in evaluating the organization's performance. Management reviews are particularly helpful in service environments because of the judgmental professional issues associated with the service sector. One organization that has developed methods for management audits and has applied this to highly diverse types of organizations with much reported success is the U.S. General Accountability Office (USGAO). The USGAO's methods are described in their audit guide referred to as the "Yellow Book" which along with examples of many of the studies can be found on their website, www.GAO.gov. Other governments, such as Canada, have also developed this type of audit and internal auditors for businesses have also developed guidelines for management audits in for profit and non-profit organizations.

1.7.11 Activity Analysis, Activity Based Costing (ABC) and Activity Based Management (ABM)

Activity Analysis: is a technique that compares the way the resource of employee time is used among similar units to establish a profile of the normal, more efficient, and less efficient units. This technique represents one powerful approach to define changes in job structure or design needed to make all units as efficient as the most efficient units. This is an area where terminology can cause unnecessary confusion. For example, program budget analysis may essentially be an activity analysis applied to an activity that is a program of the organization. An example of this would be the first year of a two-year MBA program where all students have the same first year course requirements. Here, all departments, registrar, human resources, housing, financial aid, etc. would be asked to estimate the amount of personnel time dedicated to this first year program as distinct from other programs including the second year of the MBA program. The activity analysis will provide an indication of the level of effort and human resources devoted to this program. The program budget analysis might include calculating the cost of these activities including the personnel costs resulting from the activity analysis to determining if the benefits of this program are consistent with the resources, and whether there are areas where the resources are excessive considering the objectives and the implications. Where the resources are excessive, alternative such as outsourcing food services or housing coordination or increasing fees for specific services such as housing might be explored. Here, activity analysis is key to understanding the full cost of a program.

The initial step in the activity analysis is to develop a profile of all the functions, tasks, and/or transactions of the organization or part of the organization to be evaluated. Each employee is requested to estimate the time spent on each of these activities and via tabulation of these responses and basic statistical analysis, an array of time allocations by personnel type to each of the work functions is developed. This analysis indicates where time is expended and by whom. Management evaluates these results in light of the organization's objectives, the importance of each function, and the individuals most appropriate for completing those functions. Analysis of these data suggests areas where too little or too much time is devoted to certain tasks with respect to the importance of that task. In addition, it can identify reallocation of tasks and changes of procedures that will more closely align activities with those individuals most qualified to perform them. Additionally, the shift will emphasize those activities that most benefit the organization.

In multiple office operations, activity analysis offers a view of how offices differ in the way they use personnel. These differences analyzed in light of other performance indicators may suggest patterns that are more effective, lower cost, and best practices. These patterns may then be used to adjust the way the higher cost offices provide services.

Activity analysis relies on the ability of individuals to estimate how their time is currently allocated to the designated activities. The accuracy of these data is difficult to determine and naturally must be considered in analyzing the results. The array of time spent on functions within and without each individual's and each department's activities provides insight that can lead to real productivity improvements that may not be apparent with other techniques. The assessment of which operating procedures and behavior need to be altered is, however, heavily dependent on the judgment of management rather than objective criteria (Schroeder, 1985).

Activity Based Costing – Activity Based Management: In the early 1990s, a technique referred to as activity based costing (ABC) was developed. ABC allocates all costs associated with an activity to get a full cost measure of those activities. Activity based costing focused on costs of service departments (human resources, engineering, data processing, etc.) and other costs such as energy, administration, and facilities, not easily traced to the products or services. Each department is asked to define the services and activities they provide and what drives the costs of their department. For example, the human resource department might define activities as processing new hires, responding to health benefits questions, and compliance with employee reporting regulations. They would estimate the amount of their department expenses and personnel time are associated with each of the activities (three in this human resource example). The drivers of each or the activities would be defined, such as number of new hires for the new hire activity, and type of calls for the inquiry activity, suggesting that some inquiries are more costly than other types of inquiries.

The ABC process begins by addressing similar questions raised in activity analysis, requiring each department to identify key activities and estimate the amount of resources related to these key activities that support services, products or other administrative departments. The amount of resources is then converted to dollar costs to estimate the amount of these costs associated with products or services. The result is a more precise allocation of all these indirect or less traceable costs to the products or services than was available with more crude allocation methods used in many accounting systems. The activity costs are allocated based on the activity that drives costs – referred to as cost drivers. For example, personnel costs might have been allocated the basis of number of employees in a department in a traditional accounting system. ABC might note that costs are not a function of the number of employees, but that these costs are driven by the number of inquiries, injuries, employee turnover (hires and fires), and complex employment contracts found in each department. The personnel costs associated with activities such as inquiries would be estimated and allocated to each department based on the number of actual inquiries generated during a period. The result is a more precise cost estimate than is available in many accounting systems not employing ABC. This method allocates costs common to different activities, services or products based on the activity that drives or causes these costs. In the human resource (HR) department example, management might determine that some departments use much more services than others because they have more injuries, turnover, or less training. In this case, management can assess whether the higher use is justified and valuable. If not, it will suggest an area where streamlining the process can reduce HR costs without sacrificing services that a valuable to the business.

The ABC costs can be evaluated by management in comparison to the value of these activities. For example, management may determine and eliminate excessive human resource time that was devoted to activities that add little value, such as handling minor health insurance claims. This may trigger exploration of alternative methods such as streamlining internal procedures, use of outside organizations to process claims, or changes in the arrangement with health insurers.

Another example of the way management may alter operations using ABC, which gives rise to the title of activity based management ABM, is the way one firm changed their purchase order procedures. An ABC analysis determined that after allocating all activity costs to purchase orders, the cost per purchase order was about $300. While it was well understood that every purchase order did not cost $300, on average, the cost per purchase order was $300 and a reduction in purchase orders would reduce the time and other costs associated with this activity. They adjusted their internal procedures to authorize department to make purchased below $300 without requiring purchase orders. Management concluded the control benefits over small purchases were not significant enough to warrant a cost equal to the purchase. In addition, they incorporated this analysis in their accounting system to allow them to track the cost of purchase orders over time to monitor whether the costs were increasing or decreasing. Note that ABC and ABM are a type of best practice methodology;

they attempt to locate and strive for an efficient cost, but they do not suggest that the result is efficient cost.

Consulting firms that conducted activity analysis (as distinguished from activity based costing) would be focusing on what resources are used to support a program, a department, or to produce a product or services. Activity based costing focusses on similar issues but attaches costs to these resources so they could be summed to develop costs of activities. Activity analysis was completed primarily to assess whether the amount are reasonable or seem excessive considering the objectives and value of the activity. Using activity based costing (ABC) to assess of the relative value of activities and identify ways to improve operations by lowering costs is referred to as activity based management (ABM). One key advantage of ABM is it coverts the activities to costs making it a system that can be incorporated in the accounting system of a business and which can easily be translated into earnings and ROI implications.

Each of these methods have a different advantages and limitations, and share the objective of assessing how resources that are used to advance multiple objectives in an effort to develop more efficient operations and boost profitability. They help analyze how resources are used different programs, products, and services to identify ways to provide products and services with fewer resources while maintaining quality standards.

1.7.12 Process Analysis

Process analysis or process flow analysis encompasses the review of work procedures, methods, or systems used to provide services. This entails detailed review of each procedure required to provide each service as well as the development of flow charts or other diagrammatic materials that enable the analyst to review the process. Alternative procedures, alternative design of the job function, and alternative layout of the service facilities are considered to seek ways to streamline the process to make it more effective in meeting service objectives and in reducing resource requirements (Schroeder (1985)).

1.7.13 Staffing Models

Staffing models are widely used and are based on a set of analyses and assumptions using the methodologies described in this chapter as well as other approaches such as queuing models. The staffing model reflects the amount and level or functional expertise of personnel needed for a particular type of service provided at various volume levels of activity. This approach essentially represents a standard for the human resource requirements of the service unit. Questions that warrant management scrutiny are how staffing model was developed and does it provide a standard that, if achieved, ensures high productivity? If this model reflects a type of efficient or engineered standard, it would be a useful model by which to allocate human resources to the service unit.

If the model is based on historical staffing levels and/or is not able to fully account for the mix of services provided and mix of resources used, then it will not ensure high productivity. Rather, it will promote homogeneity among service units and in some cases may result in under-allocation of resources, which can impact other dimensions such as quality.

Staffing models enable managers to quickly assess the need for resources based on a projected level of activity. It also allows comparison of actual staffing levels with the model to determine where excess resources are being utilized. Different models are needed for each type of service organization. Consequently, such a model would tend to be developed primarily for services where there are multiple offices or applications that justify the cost of developing the model. As noted, this does not represent a distinct methodology but reflects use of other methodologies. In practice, the underlying assumptions and methodology need to be reviewed to ensure that they continue to be relevant as the organization objective and operations change. While it is hard to generalize, staffing models have been used without revision past their point of relevance. Management may not be aware of or recall the assumptions behind the model and the extent to which they continue to be reasonable assumptions. As one considers which techniques make sense to manage their individual service business, another question to evaluate on a regular basis is what are the assumptions behind any staffing model in use?

1.7.14 Balanced Scorecards (BSC)

The balanced scorecard has been adopted by many businesses, government and non-profit organizations and is focused on managing the performance of these organizations. This is not a productivity management technique, but we believe its wide adoption makes it useful to clarify how the BSC relates to the methods discussed in this chapter and to DEA.

The BSC converts or redefines the strategy of the business into to a set of performance criteria and related targets. Progress in achieving strategic objectives is measured by comparing the actual results on each of the performance criteria against the target and this forms the basis for evaluating areas that require management efforts to remedy shortfalls and areas where the organization achieving its objectives. The performance measures are divided into four key areas: 1) financial, 2) customer, 3) internal business processes, and 4) learning and growth. Individuals in a business and business units have scorecards indicating the measures applied to their performance and the reward system is directly tied to the performance against the individual or department BSC targets.

For a corporation, the common and overriding goal is the financial goal related to the shareholders and this, as described earlier, is focused on achieving outstanding ROI over a long period of time. The other three elements are needed to support and generate the high ROI, including maintaining and developing profitable customers, developing new products and services that are innovative and attractive to customers, and developing the human capital to continue this cycle. Non-profit and government organizations do not share the ROI objective and need to seek other financial goals such as maintaining solvency, meeting budgets, minimizing operating costs, and in some cases, maximizing profits of some of their operations. For these organizations, the BSC highlights the fact that financial is but one of several sets of key areas of concern.

The productivity management methodologies covered in this chapter are not replacements or substitutes for the BSC. Rather, they are the techniques that would be used to boost financial results by reducing operating costs and to identify best practice internal processes to improve efficiency and reduce costs. For activities that are analyzed using productivity management methodologies, development of best practice targets may be derived from these analyses. Meaningful and objective targets are critical to the BSC, as meeting a target that is not efficient can provide a false sense of achievement.

For example, DEA has been applied to identify ways to manage physician practice patterns. For a health plan, the benefits of this can be to reduce the cost of care, which would contribute to the improved financial performance by improving the internal business processes i.e. helping physicians provide high quality care with fewer resources. In one health plan, it was noted that they are very concerned about managing costs, as cost increases result in the need to boost prices in a competitive and very sensitive market. At the same time, the plan identified another strategic concern that might be categorized in the customer satisfaction quadrant of the BSC -- concern about under-care. The health plan was concerned that pressure on physicians to keep costs down may result in some providing too little care, too few tests, too little preventive measures or immunizations etc. While using DEA to advance the financial objectives, they noticed the ability to use other information from DEA to deal with this customer quality issue. DEA identifies physicians that have low cost and low quality. The health plan is focusing on these physicians to determine if they are achieving low cost by providing under-care. This has the potential to improve the customer satisfaction, to help management learn about the practices of their service providers and to improve their business processes.

Productivity management methodologies generate insights into best practices and efficient standards, which can be used to set targets for related BSC performance goals.

Readers involved with service businesses and familiar with the balanced scorecard are likely to think of other ways the productivity methodologies discussed in this chapter can help an organization define and meet its BSC targets.

1.8. CONCLUSION

A wide range of approaches to manage productivity in a service organization have been noted in this chapter. For any one service activity, some of the techniques described may not be appropriate. Other services may require use of a combination or blending of techniques to achieve improved productivity. We believe most service organizations can enhance their productivity through applications of some of these techniques and that there are not other methods that we were able to identify as of the publication date of this book. The balance of this book focuses on the application of DEA to services.

Chapter 2

DATA ENVELOPMENT ANALYSIS EXPLAINED

The purpose of chapters 2 and 3 is to provide a clear explanation for managers who may have no background in linear programming about i) what DEA is, ii) how to apply DEA to identify paths to improve service performance, iii) how the reader can try DEA using Microsoft® Excel, and iv) how to use the Excel Solver based DEA software - DEAFrontier.

2.1. INTRODUCTION

Data Envelopment Analysis (DEA) is a very powerful service management and benchmarking technique originally developed by Charnes, Cooper and Rhodes (1978) to evaluate nonprofit and public sector organizations. DEA has since been proven to locate ways to improve service not visible with other techniques. Yet there is an anomaly surrounding this developing methodology. One of the largest US banks located over $100 million of excess annual personnel and operating costs, enough to affect their earnings per share and these savings were not identifiable with other techniques in use. While other banks have also realized improved profits through initiatives driven by DEA, we could not locate more than 10 banks in this category. While businesses have no obligation to report their internal methods, DEA has not been widely adopted by banks. Why is DEA, a method that can generate new paths to improved profits not used when other less powerful techniques continue in use? We believe that greater adoption of DEA will only be possible when it is more accessible, a key objective of this chapter and this volume. Moreover, every service organization can benefit from DEA in different ways and DEA can be adapted to help improve service productivity. Increased use by service managers will identify new strengths and benefits that can be derived from DEA along with gaps and weaknesses. The latter can set the agenda for future research on adapting DEA and will help identify areas where this methodology is inappropriate and ineffective, allowing managers to identify these types of applications of DEA.

Linear programming is the underlying methodology that makes DEA particularly powerful compared with alternative productivity management tools. DEA has been widely studied, used and analyzed by academics that understand linear programming.

Managers have not widely adopted DEA to improve organization performance, in part, because most DEA publications are in academic journals or books requiring the ability to understand linear programming and supporting mathematical notation. In fact, some managers trying to use DEA based on their understanding of academic publications have misunderstood the way to apply DEA. They erroneously attribute weak results to the technique when the problem is often due to the misapplication of DEA.

This chapter explains what DEA does, how DEA evaluates efficiency, how DEA identifies paths to improve efficiency, limitations of DEA, and how to use DEA. This will enable managers to explore and assess the value of using DEA in their service operations.

What does DEA do?

1. DEA compares service units considering all resources used and services provided, and identifies the most efficient units or best practice units (branches, departments, individuals) and the inefficient units in which real efficiency improvements are possible. This is achieved by comparing the mix and volume of services provided and the resources used by each unit compared with those of all the other units. In short, DEA is a very powerful benchmarking technique.
2. DEA calculates the amount and type of cost and resource savings that can be achieved by making each inefficient unit as efficient as the most efficient – best practice – units.

3. Specific changes in the inefficient service units are identified, which management can implement to achieve potential savings located with DEA. These changes would make the efficient units performance approach the best practice unit performance. In addition, DEA estimates the amount of additional service an inefficient unit can provide without the need to use additional resources.
4. Management receives information about performance of service units that can be used to help transfer system and managerial expertise from better-managed, relatively efficient units to the inefficient ones. This has resulted in improving the productivity of the inefficient units, reducing operating costs and increasing profitability.

The above four types of DEA information prove extremely valuable because they identify relationships not identifiable with alternative techniques that are commonly used in service organizations. As a result, improvements to operations extend beyond any performance improvements management may have achieved using other techniques.

To appreciate the power and limitations of DEA in improving efficiency, a few basic examples are described to establish various ways of defining efficiency. We believe these basic examples can be helpful for one inexperienced with DEA and interested in getting a sense of what DEA can and cannot do. If you are already familiar with basic DEA concepts, you may want to go to section 2.4.2 where the basic DEA model is presented and explained and section 2.4.3 where instructions to run DEA on Microsoft® Excel are provided.

2.2. BASIC EFFICIENCY CONCEPTS

Efficiency can be simply defined as the ratio of output to input. More output per unit of input reflects relatively greater efficiency. If the greatest possible output per unit of input is achieved, a state of *absolute* or *optimum efficiency* has been achieved and it is not possible to become more efficient without new technology or other changes in the production process.

Technical and Scale Efficiency

For example, in operating an automobile, we might measure ratios such as cost per mile or miles per gallon (MPG). We can determine the efficiency in terms of miles per gallon of auto A with its particular engine properly tuned by reference to the engineering specifications. Assume that auto A actually operates at 15 miles per gallon and its efficient miles per gallon is 20 miles per gallon. We can conclude that auto A is inefficient and that it is operating at 75% efficiency (15 MPG/20 MPG). This represents *technical* inefficiency in that excess resources are used - gasoline - to produce a unit of output-one mile of travel. Auto A should be able to increase its miles covered per gallon to 100/75 or 133% of its current miles covered (or decrease the fuel used to cover one mile by 25%). Note that regardless of the price of gas, improving the technical efficiency would reduce gas costs, possibly as much as 25%. Also, note that as long as this is the technology used, this auto could not travel more than 20 miles per gallon, so that if the auto covers 20 MPG, it is at maximum efficiency of 100% and it is not possible to exceed 100% efficiency for the given technology.

Even in this simple case, we might raise questions about conditions that influence fuel usage like road conditions, level vs. uphill travel, and the quality of the gasoline before concluding that the engine needs tuning. We will assume that these valid concerns are not issues for this introductory discussion.

If auto A actually traveled at 20 MPG, we could consider it to be 100% efficient. Auto B, designed to operate at 30 miles per gallon, which actually travels at 25 MPG, is operating at 25/30 or about 83% efficiency. If we compare actual performance, however, A is seen as less efficient than B (20/25 or 80% as efficient as B). In this case, we know that A cannot become as efficient as B with the technology and production methods used in A. This may be due to the size of A - i.e., it may be a heavier auto, which would mean that scale issues are a factor - or B may use a more advanced technology. Auto A would have to either change technologies or reduce its size to achieve higher MPG efficiency. This comparison of A and B does not allow us to observe that B is also *inefficient* with respect to the technology it employs. In order to gain some insights about B's inefficiency, it is necessary to compare autos of the same type that have better tuned engines and therefore operate at more than 25 MPG.

Key Point 1: *When we focus on service organizations, we generally cannot determine what the engineered, optimum or absolute efficient output-to-input ratio is. This is in contrast to the auto example where it was possible to determine the efficient engine performance. Consequently we cannot determine whether a service unit is absolutely efficient. We can, however, compare several service unit output-to-input ratios and determine that one unit is more or less efficient than another - benchmarking. The difference in efficiency will be due to the technology or production process used, how well that process is managed, and/or the scale or size of the unit.*

Price Efficiency

We might choose to use cost per mile instead of the MPG measure of performance. One auto of type A, (A1), might be operating at 20 MPG using gasoline that costs $1 per gallon, resulting in a cost of 5 cents per mile. Another auto of type A, (A2), might travel 20 MPG but pay 80 cents per gallon, resulting in a cost of 4 cents per mile. A1 is then only 4/5 or 80% as efficient as A2, but this is not due to differences in the way the engine is tuned or the technology employed. Rather it is due to *price* efficiency; A2 buys its fuel (input) at a lower cost. Again we don't know if A2 is absolutely efficient because it may be possible to buy gasoline at a price below 80 cents per gallon. Nevertheless, we know that A1 is inefficient and could do better in terms of price efficiency. However, if a manager misinterpreted the higher cost of A1 and requested the mechanics tune the engine to improve the performance and lower cost per mile, it would just waste the cost of the tune-up, as the engine is already 100% efficient. The cost per mile, similar to the information on operating expenses in an income statement, provides no indication of the reason the costs are higher or lower than other units or past periods. DEA allows a manager to look behind the accounting information to separate excess costs due to technical and scale efficiency from price efficiency to understand what type of actions they can initiate to reduce cost and improve profitability.

We realize the price of gasoline can be three times the hypothetical in the US and that the gallon cost might be replaced by the cost per liter in Europe. We note this to assure the reader that the intent of the book is to provide usable practical guidance in applying DEA and that the simple illustrations should be accepted as a means to learning how DEA works. Actual applications with real data that yielded substantial verifiable benefits are provided in the latter part of this book beginning with Chapter 6.

Key Point 2: *If inputs and/or outputs are measured in dollars rather than physical units, the efficiency differences we observe can be due to price efficiency as well as scale and technical efficiency. If we use both methods, physical and dollar measures, we can begin to segregate price from technical and scale efficiency. To understand which units are inefficient and how to improve them, we need to separately measure all the types of inefficiency present. DEA is one method of separating technical and scale efficiency from price efficiency, enabling it to locate methods of improving profitability of service organizations that already appear relatively profitable based on accounting measures reflected in an income statement.*

Allocative Efficiency

Consider two employees, John and Mary, who each own two autos, one of type A and one of type B, and they can use either one or both for business transportation. If they use their respective A and B cars in different proportions, they may have different travel costs. Assume both A autos get 20 MPG, both B autos get 25 MPG, and all gas is purchased at $1 per gallon, so that the As cost 5 cents per mile and the Bs cost 4 cents per mile. If John uses only B and Mary uses A and B equally, John's cost is 4 cent per mile while Mary's cost is 4.5 cents per mile. Comparing these travel cost ratios, we could conclude that Mary is less efficient than John. His costs are 4.5/5 or 90% of Mary's and Mary is thus only 90% as efficient as John. This is an example of *allocative* inefficiency, which results from an inefficient mix of inputs - cars - used to produce the output - miles traveled. Note that both cars A and B are equally efficient with respect to MPG and input price. Note that a manager trying to reduce travel costs without the ability to segregate different types of inefficiency could erroneously suggest that Mary's cars need tune-ups or that she should try to find lower price gas, while Mary and John are already using efficiently tuned autos and paying the same price for gas. One way to reduce costs is to change the mix of use of Auto A versus Auto B. This conclusion requires the ability to segregate allocative efficiency from price and technical efficiency. We will illustrate ways DEA can assist in this type of analysis in the subsequent chapter.

Key Point 3: When more than one input and/or output are involved in the production process, inefficiencies can also be due to the mix of inputs used to produce the mix of outputs, which is referred to as allocative efficiency.

2.3. RELATIVE WEIGHTS AND COSTS FOR INPUTS AND OUTPUTS

Consider another scenario represented in Table 2-1. John and Mary use each of their cars to cover the miles needed to accomplish their respective assignments. Number of miles traveled is the output. There are two inputs: usage of car A and usage of car B. This car input could be measured in dollars costs or in amount of input units such as fuel, hours of use, or even miles.

Table 2-1. Car Usage Example with Price and Allocative Inefficiencies

			Actual Inputs			Actual Output
Car	# miles traveled	MPG	Cost per gallon	Total Cost	Cost per mile	# of miles
			John			
A	80	20	1.00	\$4.00	0.05	100
B	20	25	1.00	\$0.80	0.04	Miles
Total	100			\$4.80	0.048	
			Mary			
A	40	20	1.50	\$3.00	0.075	200
B	160	25	1.50	\$9.60	0.06	Miles
Total	200			\$12.60	0.063	

How can these two employees' efficiency with respect to the use of automobiles be evaluated? One common approach is cost per mile. This would suggest that John is more efficient than Mary as his cost was 4.8 cents per mile versus 6.3 cents per mile for Mary. This reflects the cost and amount of inputs used. From Table 2-1, we can observe that cost per mile is influenced by the cost per gallon. Indeed, John is more efficient than Mary because he pays less per gallon. This ratio analysis could be concluded at this point; and Mary could simply be told to use John's gas supplier or one with comparable costs. This would be a sub-optimal result, however, because it overlooks the car utilization; i.e., price efficiency is dealt with but we have not considered whether other efficiency improvements are possible such as changing mix of type A and type B automobiles used. Another level of ratio analysis would be to compare the physical inputs with the outputs: there are two inputs, car A and B; how much of these are used? Since a ratio can include only one numerator and one denominator, the simplest approach is to add up the inputs and outputs to get one input and output number.

John's inputs are 20 miles of A and 80 miles of B totaling 100 car units or miles of input. This yields an output-to-input ratio of 100/100 = 1. Mary's corresponding output-to-input ratio is 200/(160 + 40) = 1. This ratio makes John and Mary appear equally efficient even though we can observe that this is not an accurate conclusion.

The problem is that this ratio does not reflect the different usage levels of cars A and B by John and Mary. Note, however, that if Mary does finally buy gas from John's supplier, her cost per mile will be 2/3 its current level, or 4.16 cents per mile, which is lower than John's cost of 4.8 cents per mile. Mary will then appear more efficient than John. There remain inefficiencies that are not observable using both the cost-per-output ratio and combined the physical input-to-output ratio. In this example we can see that Mary used more of car B, which had higher MPG than car A. Mary was more efficient because of the *mix* of physical inputs used.

To make the efficiency ratios sensitive to the input and output mix, we would have to weight the inputs by their relative values. For example, we know that B uses 80% of the gas that A uses and therefore the cost per unit of B is 80% of A.

Applying this relative value weight to A and B usage, we calculate adjusted input to output efficiency ratios for John: 100/([80][1] + 20 [0.8]) = 100/96 = 1.04 and for Mary: 200/([40][1] + [160][0.8]) = 200/168 = 1.19

Only after the inputs are weighted by relative values and costs, does the ratio reflect Mary as more efficient. She gets 1.19 miles per weighted car input unit while John only gets 1.04 mile per weighted car input unit.

The relative weights needed to value inputs (and outputs) are often not available. This is particularly true for service organizations. Without these weights, ratio analysis may be only marginally helpful and possibly misleading in multiple-output, multiple-input applications. This inability to identify reliable relative weights for different inputs and outputs limits the ability to use operating ratios to gain insights into ways to manage and improve performance. DEA has the ability to analyze relative performance when such weights are not available making it particularly effective in service environments where these weights are not available. This attribute – the ability to incorporate multiple inputs and outputs in their natural units without knowledge of the relative weights – makes DEA uniquely suited for evaluating many service organizations and providers.

Key Point 4: Ratios can provide very useful managerial information about efficiency; however, they are incapable of accommodating multiple inputs and outputs when accurate objective relative weights for inputs and outputs are not known.

We noted earlier that car A could expect to operate at 30 MPG if it were properly tuned to its engineered efficiency level. The above analysis yields no hint of this type of inefficiency – whether John and/or Mary would be more efficient by having their cars tuned. This would require that we have an efficient standard to compare with actual results for a complete evaluation. The efficient standard is generally not available in service environments. Consequently, there will often be possibilities for efficiency improvements that will not be apparent from available analytic techniques including DEA.

Finally, in proceeding with the DEA discussion, note that the easily observed causes of inefficiency in this simple example will not be readily observable in service units where several inputs and outputs are involved and where there may be 20 or 200 or 2,000 service units or providers being evaluated instead of two. DEA addresses many but not all of these problems and it has the ability to complement common analytical tools like ratio analysis to find ways of improving performance that are otherwise invisible as will be illustrated in the balance of this chapter.

2.4. DATA ENVELOPMENT ANALYSIS

2.4.1 How DEA Works and How to Interpret the Results

We now illustrate how DEA is used to evaluate efficiency by means of the simplified bank branch example noted in Table 2-2. This analysis assumes only one type of transaction (one service such as check cashing or receiving deposits) and two types of resources used to process these transactions - bank teller hours (H) and supply dollars (S). This example was selected because it lends itself to graphic description, and because it is simple enough to be analyzed without DEA. Hence, the results can be compared to an independent analysis of efficiency. Note that DEA is most valuable in complex situations where there are multiple outputs and inputs, which cannot be readily analyzed with other techniques like ratios, and where the number of service organization units being evaluated is so numerous that management cannot afford to evaluate each unit in depth.

Table 2-2. Illustrative Example of Five Bank Branches

	Service Output	Input Used	
Service Unit	Transactions Processed (T)	Teller Hours (H)	Supply Dollars (S)
B1	1,000	20	300
B2	1,000	30	200
B3	1,000	40	100
B4	1,000	20	200
B5	1,000	10	400

Assume that there are five bank branches (B1, B2, B3, B4, and B5) and that each processes 1,000 transactions (e.g., deposits) during one common time period (e.g., week, month, year) by jointly using two inputs: tellers measured in labor hours (H) and supplies measured in dollars (S). (See Table 2-2 for a summary of the outputs and inputs.)

The problem facing the manager is to identify which of these branches are inefficient and the magnitude of the inefficiency. This information can be used to locate the branches that require remedial management action, to reward the more efficient managers, and/or to determine the management techniques used in the more efficient branches that should be introduced into less efficient branches. While the manager can observe the number of transactions processed and the amount of resources (H and S) used, he or she does not know the efficient output-to-input relationship. That is, the efficient amount of labor and supplies needed for each transaction is not readily determinable. The problem is illustrated in Figure 2-1.

In this example, it can be observed that B1 and B2 are relatively inefficient. B1 produced the same output level as B4 but used 100 more supply dollars (S) than B4. B2 also produced the same output level as B4 but achieved this by using 10 more teller labor hours. With the information available in Table 2-2, it is not possible to determine whether B3, B4, or B5 is more or less efficient. While information about relative prices might allow one to rank B3, B4 and B5, the finding that B1 and B2 are inefficient would not change. That is, B1 and B2 should be able to reduce inputs without reducing outputs regardless of the price of the inputs.

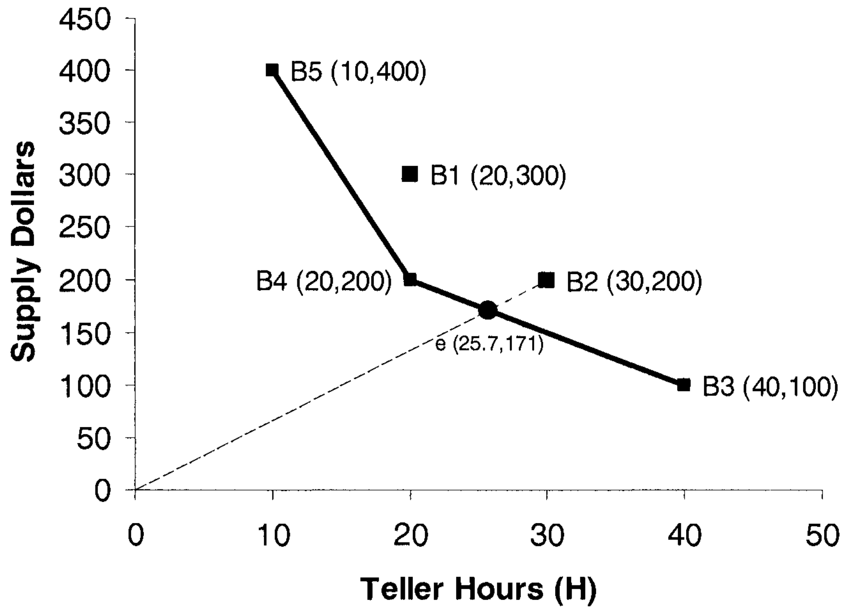


Figure 2-1. Graphic Representation of the Five Bank Branches

DEA compares each service unit with all other service units, and identifies those units that are operating inefficiently compared with other units' actual operating results. It accomplishes this by locating the best practice or relatively efficient units (units that are not less efficient than other units being evaluated). It also measures the magnitude of inefficiency of the inefficient units compared to the best practice units. The best practice units are relatively efficient and are identified by a DEA efficiency rating of $\theta = 1$. The inefficient units are identified by an efficiency rating of less than 1 ($\theta < 1$)[1]. DEA will provides an efficiency rating that is generally denominated between zero and 1, which will interchangeably be referred to as an efficiency percentage between the range of zero and 100%. The upper limit is set as 1 or 100% to reflect the view that a unit cannot be more than 100% efficient. Models in chapter 4 will describe a type of DEA analysis that allows the upper limit to exceed 1 or 100% for particular applications. Hence the computer DEA output may indicate the unit

[1] The symbol θ is used to denote the efficiency measure (score) consistent with the original DEA literature.

efficiency $\theta = 0.43$, which may also be referred to as a 43% efficiency rating.

Table 2-3. DEA Results for Five Bank Branches

Service Unit	Efficiency Rating (θ)	Efficiency Reference Set (ERS)	
B1	85.7%	B4 (0.7143)	B5 (0.2857)
B2	85.7%	B3 (0.2857)	B4 (0.7143)
B3	100%		
B4	100%		
B5	100%		

Table 2-3 reports the results of DEA applied to these five branches. DEA identified the same inefficient branches that were identifiable through observation of the data. B1 and B2 have efficiency ratings below 100%, which identifies them as inefficient. In addition, DEA focuses the manager's attention on a subgroup of the bank branches referred to as the efficiency reference set in Table 2-3. This efficiency reference set (ERS) includes the group of service units against which each inefficient branch was found to be most directly inefficient. For example, B1 was found to have operating inefficiencies in direct comparison to B4 and B5. The value in parentheses in Table 2-3 represents the relative weight assigned to each efficiency reference set (ERS) member to calculate the efficiency rating (θ). Figure 2-1 illustrates this using B2 as an example. (If a service unit's efficiency rating is 100%, then this unit is its own ERS and we generally do not report it as an ERS, which is the reason B3, B4, and B5have not reported ERS in the Table 2-3.)

DEA has determined that, among the five bank branches, B5, B4, and B3 are relatively efficient. In this simple case, the solid line in Figure 2-1, which locates the units that used the least amount of inputs to produce their output level, represents this. These three branches, B5, B4 and B3 comprise the best practice set or best practice frontier. No indication is provided as to which if any of these three is more or less efficient than the other two. As noted earlier, all three could be somewhat inefficient. The best practice units are those that are not clearly inefficient compared with other units being evaluated.

DEA indicates that B2 is inefficient compared to point e on the line connecting B4 and B3 in Figure 2-1. One way for B2 to become efficient is for it to reduce its inputs to 85.7% of its current level. This would move B2 onto the relatively efficient production segment at point e in Figure 2-1, which reflects the use of 25.7 teller hours (0.857 x 200) and use of 171 supply dollars (0.857x200). DEA provides information to complete the calculation suggested in Figure 2-1. This is illustrated in Table 2-4. (In Table 2-4, the inputs and outputs of the ERS branches B3 and B4 are multiplied by the weights derived by DEA noted in Table 2-1, 0.2857 and 0.7143, respectively. These are then added together to create a composite branch that provides as much or more services as the inefficient branch, B2, while also using less inputs that B2. These ERS weights are generally referred to as Lambda - λ - values in the DEA models described below.)

Table 2-4. Inefficiency in Branch B2 Calculated by DEA

Outputs		Outputs & Inputs of B3			Outputs & Inputs of B4		Efficiency Reference Set for Service Unit B2
Transaction Processed(T)		1,000			1,000		1,000
Inputs	(0.2857)×		+	(0.7143)×		=	
Teller Hours (H)		40			20		25.7
Supply $ (S)		100			200		171

The composite for B2 can then be compared with the inefficient unit B2 as follows:

	Column 1 Composite Outputs & Inputs (from above)	**Column 2** Branch B2 Actual Outputs & Inputs	Column 2 – Column 1	
(T)	1,000	1,000	0	
(H)	25.7	30	4.3	Excess Inputs Used by Branch B2
(S)	171	200	29	

Table 2-4 indicates that a mixture of the operating techniques utilized by B3 and B4 would result in a composite hypothetical branch that processes the same number of transactions (1,000) as B2 but that requires fewer inputs than B2. Hence, by adopting a mixture of the actual techniques used by B3 and B4, B2 should be able to reduce teller hours by 4.3 units and supply dollars by 29 units without reducing its output level. A similar calculation can be completed for each inefficient unit located by DEA analysis. These potential savings located with DEA, not identifiable with other available techniques, have been converted into substantial real savings for service organizations.

At this point it must be re-emphasized that DEA results are most useful when there are multiple outputs and inputs, and where the type of intuitive analysis that could be applied to verify the DEA results in the above example would not be possible. Nevertheless, the efficiency rating, the efficiency reference set, the analysis performed in Table 2-4, and the ability to determine alternative paths that would make an inefficient unit efficient would all be readily available to management. Applications to numerous organizations suggest that the representation in Table 2-4 is one of the more direct ways to summarize and explain what DEA has achieved and its implications for management. The manager can see that a combination of existing branches results in a more efficient use of resources while providing the same services (outputs) as the inefficient branch. This can be adequate evidence to cause the manager to question why can't the inefficient branch 2 be as efficient as branches 3 and 4. While the weights applied to branch 3 and 4 to get the composite directly come from the DEA linear programming formulation, the manager does not need to know how those weights were developed and how DEA determined that branch 3 and 4 were identified as ERS branches that should be compared with branch 2.

In summary, the interpretation of DEA results tends to proceed in the following order

- The efficiency ratings are generated as in Table 2-4. Units that are efficient ($\theta = 1$) are relatively, and not strictly, efficient. That is, no other unit is clearly operating more efficiently than these units, but it is possible that all units, including these relatively efficient units, can be operated more efficiently. Therefore, the efficient branches (B3, B4, and B5) represent the best existing (but not necessarily the best possible) management practice with respect to efficiency.

- Inefficient units are identified by an efficiency rating of $\theta < 1$. These units (B1 and B2) are strictly inefficient compared to all other units and are candidates for remedial action by management. In fact, the inefficiency identified with DEA will tend to understate, rather than overstate, the inefficiency present because of the nature of linear programming which seeks to maximize the efficiency rating. .

- The efficiency reference set (ERS) indicates the relatively efficient units against which the inefficient units were most clearly determined to be inefficient. The presentation in Table 2-3 summarizes the magnitude of the identified inefficiencies by comparing the inefficient unit with its efficiency reference set.

- The results in Table 2-4 indicate the following: B2 has been found to be relatively less efficient than a composite of the actual output and input levels of B3 and B4. If a combination of the operating techniques used in B3 and B4 were utilized by inefficient B2, B2 should be able to reduce the number of hours used by 4.3 units and the amount of supplies used by 29 units while providing the same level of services. Of course, management can also use DEA to identify other methods or combinations of methods to improve the efficiency of inefficient units.

2.4.2 The Mathematical Formulations of DEA

The linear programming technique is used to find the set of coefficients (*u*'s and *v*'s) that will give the highest possible efficiency ratio of outputs to inputs for the service unit being evaluated.

Table 2-5 provides a DEA mathematical model. In the model,
j = number of service units (SU) being compared in the DEA analysis
SUj = service unit number j
θ = efficiency rating of the service unit being evaluated by DEA
y_{rj} = amount of output r used by service unit j
x_{ij} = amount of input i used by service unit j
i = number of inputs used by the SUs
r = number of outputs generated by the SUs
u_r = coefficient or weight assigned by DEA to output r
v_i = coefficient or weight assigned by DEA to input i

The data required to apply DEA are the actual observed outputs produced y_{rj} and the actual inputs used x_{ij}, during one time period for each service unit in the set of units being evaluated. Hence, x_{ij} is the observed amount of the *i*th input used by the *j*th service unit, and y_{rj} is the amount of *r*th output produced by the *j*th service unit.

If the value of θ for the service unit being evaluated is less than 100%, then that unit is inefficient, and there is the potential for that unit to produce the same level of outputs with fewer inputs. The theoretical development of this approach is discussed in detail in Cooper, Seiford and Tone (2000) and Zhu (2003). Rather than reproduce this discussion, DEA will be explained with several simple applications and with emphasis on how to apply it, how to interpret the results and the implications for managing productivity.

Table 2-5. DEA Mathematical Model

Objective Function

$$\text{Maximize } \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \ldots + u_r y_{ro}}{v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo}} = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

(Maximize the efficiency rating θ for service unit o.)

This is subject to the constraint that when the same set of u and v coefficients is applied to all other service units being compared, no service unit (SU) will be more than 100% efficient as follows:

$$\text{SU1} \quad \frac{u_1 y_{11} + u_2 y_{21} + \ldots + u_r y_{r1}}{v_1 x_{11} + v_2 x_{21} + \ldots + v_m x_{m1}} = \frac{\sum_{r=1}^{s} u_r y_{r1}}{\sum_{i=1}^{m} v_i x_{i1}} \leq 1$$

$$\text{SU2} \quad \frac{u_1 y_{12} + u_2 y_{22} + \ldots + u_r y_{r2}}{v_1 x_{12} + v_2 x_{22} + \ldots + v_m x_{m2}} = \frac{\sum_{r=1}^{s} u_r y_{r2}}{\sum_{i=1}^{m} v_i x_{i2}} \leq 1$$

...

$$\text{SUo} \quad \frac{u_1 y_{1o} + u_2 y_{2o} + \ldots + u_r y_{ro}}{v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo}} = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \leq 1$$

...

$$\text{SUj} \quad \frac{u_1 y_{1j} + u_2 y_{2j} + \ldots + u_r y_{rj}}{v_1 x_{1j} + v_2 x_{2j} + \ldots + v_m x_{mj}} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1$$

$u_1, \ldots, u_s > 0$ and $v_1, \ldots, v_m \geq 0$

DEA differs from a simple efficiency ratio in that it accommodates multiple inputs and outputs and provides significant additional information about where efficiency improvements can be achieved and the magnitude of these potential improvements. Moreover, it accomplishes this without the need to know the relative value of the outputs and inputs that were needed for ratio analysis (see Key Point 4 above).

Assume that the DEA evaluation would begin by evaluating the efficiency of bank branch B2 in Table 2-2. Based on the DEA model (Table 2-5), the problem would be structured as described below using the data in Table 2-2.

Calculate the set of values for u_1, v_1, and v_2 that will give branch B2 the highest possible efficiency rating:

$$\text{Maximize } \theta = \frac{u_1(1000)}{v_1(30) + v_2(200)}$$

This is subject to the constraint that no service unit (in this case bank branch) can be more than 100% efficient when the same values for u_1, v_1, and v_2 are applied to each unit:

B1 $$\frac{u_1(1000)}{v_1(20) + v_2(300)} \leq 1$$

B2 $$\frac{u_1(1000)}{v_1(30) + v_2(200)} \leq 1$$

B3 $$\frac{u_1(1000)}{v_1(40) + v_2(100)} \leq 1$$

B4 $$\frac{u_1(1000)}{v_1(20) + v_2(200)} \leq 1$$

B5 $$\frac{u_1(1000)}{v_1(10) + v_2(400)} \leq 1$$

DEA calculates the efficiency rating for B2 to be 85.7% and the value for v_1 = 1.429 for hours, v_2 = 0.286 for supply \$s, and u_1 = 0.0857 for service units. DEA would be rerun for each branch in the objective function as was done above to branch B2.

Management is provided with alternative paths to improve the efficiency of bank branches in Table 2-2. For example, for branch B2, one path suggested in Table 2-4 (or model (2.2)) is for B2 to reduce H by 4.3 units and to reduce S by 29 units. Other paths can also be ascertained from the model (2.1). For branch B2, we have an optimal value of u = 0.000857 for transaction outputs, v1 = 0.01429 for teller hours (H) and v2 = 0.00286 for supply dollars (S). This means that for each reduced teller hour, the efficiency of B2 increases by 1.43%. For each supply dollar decrease, the efficiency of B2 will increase by 0.286%. For B2 to become relatively efficient, it must increase its efficiency rating by 14.3 percentage points (100-85.7%). Hence, B2 can become efficient by decreasing H by 10 hours (10 hours X 1.43=14.3%) or by decreasing S by 50 units (50×0.286%=14.3%), or by some combination of these reductions in H and S. Of course, management would choose a particular set of operating changes based on an evaluation of cost, practicality, and feasibility for that particular organization unit.

Note that optimal multipliers calculated by DEA are objectively determined weights that may not correspond to relative values that a bank would assign to outputs and inputs. This is actually a strength and is *not* a weakness of DEA. A bank branch located as inefficient using DEA is so identified only after all possible weights have been considered to give that branch the highest rating possible consistent with the constraint that no branch in the data set can be more than 100% efficient. Hence, any other set of weights applied to all branches would only make an inefficient branch appear equally or less efficient; that is, *DEA gives the benefit of the doubt to each branch or service unit in calculating the efficiency value. In addition, DEA will not erroneously locate an efficient unit as inefficient.*

Managerial perspective on the DEA formulation: One of the two common forms of basic DEA is presented above and the other common form, referred to as the dual model, is presented below. These and other models are presented in detail to ensure that they are available to readers who want to understand how the model works. The manager will not need to understand these to be able to use them. However, there are a few observations that stem from the model that can be very useful for managers to understand as they use DEA.

1. DEA gives the "benefit of the doubt" to each unit being evaluated trying to make it look as efficient as possible in comparison with the other units. This means that the inefficiencies noted would tend to understate the actual inefficiencies that may be present. This is manifest by the maximization function in the DEA model above and in most DEA models. This bias makes this a tool that managers can use with confidence. When a DEA analysis is determined to be complete in terms of using appropriate inputs and outputs, it offers paths to achieve real improvements in performance. The amount of the improvements that are technically available would be at least as great as the amount identified with DEA. Indeed, the conservative nature can occasionally result in all or almost all the units being assigned an efficiency rating of one. This result is understandably disappointing to managers, as it suggests that DEA can find little opportunity to increase efficiency of the organization. At the same time, when DEA points to inefficiencies, they are real, often substantial, and generally not identifiable with other techniques.

2. The weights are assigned by DEA to make each service unit look as efficient as possible. In the above example, we did not know the relative value of teller hours versus supplies. DEA calculated a weight for these for B2 based on the linear program formulation. The weights calculated for hours was $v_1 = 1.429$ and $v_2 = 0.286$ for supply dollars ($s). These weights when plugged into the above model suggest that the ERS branches B3 and B4 have efficiency ratings of 1.0 or 100% and B2 has an efficiency rating of 0.857 or 85.7%. Roughly, this means B2 is using about 15% excess resources based on the DEA analysis compared with B3 and B4 and these savings would be achievable if B2 operated more like B3 and B4.

   When the manager of B2 is informed that their unit is not performing as well, the manager could question whether the weights used for the inputs of hours and supplies are accurate weights, particularly since these weights are not known in practice. That manager can be challenged to find any other set of weights that if applied to all the branches would make B2 look more efficient while not allowing any other branch to be more than 100% efficient. They would fail this challenge. The set of weights calculated already make B2 appear as efficient as possible compared with the other branches. Any other set of weights applied to all the units would result in B2 having an even lower efficiency rating than the current rating of 85.7%. This also means that if we substitute another set of weights that are believed to be more reflective of the market than the weights assigned by DEA, the inefficiency will be greater and the potential benefits of improving the inefficient units to approach the best practices will be greater than estimated with the model above. (Readers are encouraged to explore the impact of using a different set of weights.)

3. Relative weights: The weights assigned to the inputs and outputs have managerial and analytic value. In DEA models discussed in chapter 5, influencing these weights allows the manager to substantially increase the DEA insights about ways to improve performance.

   The weights information has also been used for analytic purposes, sometimes overestimating the information content of these variables. For example, the values $v_1 = 1.429$ for hours, $v_2 =$

0.286 for other expenses define the slope of the line between B3 and B4 in Figure 2-1. The slope is -5, which is v_1/v_2. This has been characterized as a rate of substitution suggesting that the units operating in this segment of the efficient frontier, B3,B4 could substitute $5 of supplies for one teller hour. This describes the slope of the actual line. There are some DEA user that will interpret this slope as the actual rate of substitution or the efficient rate of substitution for these inputs. If a manager were planning to use these variables to derive rates of substitution, careful testing and analysis of the results will be needed before one can rely on these data in this way. At this point, it is sufficient to caution users about misuse of this part of the DEA information. We will clarify the issues in the next Chapter when we have a more elaborate example to illustrate the potential value and misuse of these weights.

We should also note that multiple optimal weights are very likely to be present When we solve the DEA model, weights generated by computer runs can be different from each other. We do not recommend use optimal DEA weights (or multiplier) directly.

To run DEA on a standard linear program package, the fractional forms in Table 2-5 are algebraically reformulated as follows[2]:

Maximize $\theta = u_1 y_{1o} + u_2 y_{2o} + \ldots + u_r y_{ro} \; (= \sum_{r=1}^{s} u_r y_{ro})$

Subject to the constraints that[3]

$$v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo} = \sum_{i=1}^{m} v_i x_{io} = 1$$

$$u_1 y_{1j} + u_2 y_{2j} + \ldots + u_r y_{rj} \le v_1 x_{1j} + v_2 x_{2j} + \ldots + v_m x_{mj}$$

[2] This expression maximizes the numerator for the unit being evaluated, trying to assign it the highest possible productivity rating.

[3] This expression sets the denominator for the unit being evaluated equal to 1. This is related to the Charnes-Cooper transformation (Charnes and Cooper, 1961). Although we use the same (weights) notions, the weights in this model are actually different from the ones in Table 2-5. Interested reader is referred to the first chapter in Cooper, Seiford and Zhu (2004) for the transformation involved.

The above expression in standard mathematical notation is:

$$\sum_{r=1}^{s} u_r y_{rj} \le \sum_{i=1}^{m} v_i x_{ij} \ .$$

That is, the DEA model presented in Table 2-5 is actually calculated as

Maximize $\sum_{r=1}^{s} u_r y_{ro}$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, j = 1, \ldots, n \tag{2.1}$$

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

$$u_r,\ v_i \ge 0$$

where, we assume that we have n service units.

To obtain the information provided in Table 2-4, one needs to employ the dual linear program to model (2.1). That is,

$\min \theta$

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io} \qquad i = 1,2,\ldots,m; \qquad (a) \tag{2.2}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} \qquad r = 1,2,\ldots,s; \qquad (b)$$

$$\lambda_j \ge 0 \qquad j = 1,2,\ldots,n. \qquad (c)$$

The dual is seeking the efficiency rating, minimize θ, subject to the constraint (a) that the weighted sum of the inputs of the other service units is less than or equal to the inputs of the service unit being evaluated and (b) that the weighted sum of the outputs of the other service units is greater than or equal to the service unit being evaluated. The weights are the λ (lambda) values. The other service units with non-zero lambda values are the units in the efficiency reference set (ERS). For example, in Table 2.4, DEA was evaluating B2 and the λ value for B3 was 0.2857 and for B4 was 0.7143. The efficiency rating θ for B2 was 85.7%.

The result is the weighted sum of the inputs was 25.7 Teller Hours which was 85.7%, the efficiency rating of the amount used by B2 and the weighted sum of the Supply $s was 171 which is 85.7% of the amount used by B2 meeting constraint (b) above. The weighted sum of the outputs, transactions processed, was 1000, which is equal to the output of B2. When the θ is below 1 or 100%, the above situation results where there are groups of ERS units that produce as much or more outputs and use less inputs offering a path to improve efficiency of the inefficient units. When DEA tries to minimize θ for the service unit being evaluated and it cannot find λ weights that will generate an efficiency level below 1 or 100%, this defines a relatively efficient unit where there is no opportunity to improve efficiency compared with the performance of other service units in the data set.

In DEA, model (2.1) is referred to as "multiplier model" where u_r and v_i represent output and input multipliers (weights), respectively. Model (2.2) is referred to as "envelopment model". We next illustrate how to solve DEA models (2.1) and (2.2) via Microsoft® Excel Solver and how to obtain the information on the Efficiency Reference Set.

2.4.3 Solving Envelopment DEA Model as a Linear Program in Spreadsheets

This section describes the process of incorporating the DEA analysis into a Microsoft® Excel spreadsheet (Excel). This should prepare a new DEA user to apply it using Excel with or without the software program included in this volume. We believe this is also useful for two distinct reasons. First, it demystifies what is happening in other elaborate DEA codes or programs to emphasize that this is not a highly complex process and that it is understandable at the level of basic algebra. Second, the process is inputting or recording the equations into this excel program helps familiarize the user with the basic relationships-equations in the DEA model.

The programming elements described below are referenced to the basic DEA model to allow the reader to create and run a DEA program that runs on Excel. This involves a few steps that may utilize Excel capability unfamiliar to the reader, but which can readily be used to develop a working DEA program if the following steps are carefully followed. The process involves the following steps:

1. Install and enable the Excel spreadsheet to use a program that is referred to as an "add-in". The "add-in" that will be enabled for

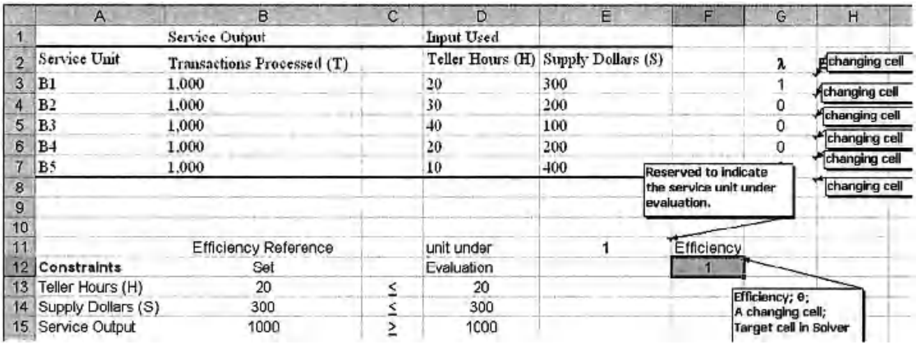
DEA is the Excel Solver. Instructions to enable the Solver to operate in your excel program are provided in Excel (the help menu item), or please refer to section 2.4.3.1.

2. The Excel solver allows you to input equations that can be maximized, minimized or just solved for. You will have instructions herein about how to input the constraints in the basic DEA models (2.1) and (2.2) into the Solver.
3. To run the Solver model and solve the DEA evaluations for a single unit, you will need to set a few parameters in the solver that indicate the type of DEA analysis being performed. Instructions on which parameters to use are provided.
4. The DEA analysis to evaluate/benchmark a single service unit compared with a set of service units can then be run. This is applied to the bank branch example described in the previous section, generating the same results attainable by hand calculations. This can be run for each of the 5 sample branch units by running the DEA program five times. To eliminate the need to rerun the Solver model each time, a task that can become tedious and costly, there are programming methods to have the program rerun itself for each service unit in the data set. This is a programming requirement that is independent of the DEA methodology and requires developing a program to rerun the analysis for every service unit in the data set.
5. The program to iteratively run the DEA analysis to evaluate the productivity and generate the full set of analytic data on every branch in the data set is provided. Incorporating this set of commands into the Excel macro will enable you to run the full DEA analysis on a set of data with one command. These commands are incorporated into the Visual Basic for Applications (VBA) module of Excel. VBA is the computer language that underlies much of the operating systems that are used in many programs including Excel, Word and Power Point. The reader does not need to be familiar with VBA, and indeed that is a very interesting topic that one could study independent of this topic. All the instructions to use VBA to enable the iterative DEA analysis are provided including access to an electronic version of these VBA commands that can just be copied and pasted into the excel program. This does not require understanding of VBA.
6. While we try to provide a clear path to completing a DEA analysis on your computer, a CD attached to this book will provide all the sample excel files and ready-to-use DEA

software *DEAFrontier*. While the CD provides a completely reliable and viable short-cut to using DEA, those interested in becoming familiar with the workings of DEA should find the following a useful way to be aware of what the DEA program is doing.

We first illustrate how to formulate the envelopment model (2.2) in a spreadsheet, and then illustrate how Excel Solver can be used to calculate the efficiency for the bank branches in Table 2-2.

We begin by organizing the data in Table 2-2 in a spreadsheet as shown in Figure 2-2. A spreadsheet model of DEA model (2.2) contains the following four major components: (1) cells for the decision variables (λ_j); (2) cell for the objective function (efficiency, θ); (3) cells containing formulas for computing the DEA efficiency reference set (the left-hand-side of the constraints) ($\sum_{j=1}^{n} \lambda_j \ x_{ij}$ and $\sum_{j=1}^{n} \lambda_j \ y_{rj}$); and (4) cells containing formulas for computing the service unit under evaluation (right-hand-side of the constraints) ($\theta \ x_{io}$ and y_{ro})



	A	B	C	D	E	F	G	H
1		Service Output		Input Used				
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)		λ	changing cell
3	B1	1,000		20	300		1	changing cell
4	B2	1,000		30	200		0	changing cell
5	B3	1,000		40	100		0	changing cell
6	B4	1,000		20	200		0	changing cell
7	B5	1,000		10	400			changing cell
8								changing cell
9								
10								
11		Efficiency Reference		unit under	1	Efficiency		
12	Constraints	Set		Evaluation		1		
13	Teller Hours (H)	20	≤	20				
14	Supply Dollars (S)	300	≤	300				
15	Service Output	1000	≥	1000				

Figure 2-2. Spreadsheet Model for DEA Model (2.2)

In Figure 2-2, cells G3 through G7 represent λ_j (j = 1, 2, ..., 7). Cell F12 represents the efficiency score θ, which is the objective function in model (2.2). Cell E11 is reserved to indicate the service unit under evaluation. Note that we will be solving for the λ_j values and the efficiency rating (cell F13). Figure 2-2 includes values of 1 and zero in these cells to illustrate the setup of the spreadsheet just before solving the DEA problem with Excel.

For the DEA efficiency reference set (left-hand-side of the model (2.2)), we enter the following formulas that calculate the weighted sums of inputs and outputs across all service units, respectively[4].

Cell B13 = SUMPRODUCT(D3:D7,$G$3:$G$7)
Cell B14 = SUMPRODUCT(E3:E7,$G$3:$G$7)
Cell B15 = SUMPRODUCT(B3:B7,$G$3:$G$7)

In Figure 2-2, the above equations are embedded in the cells as with any Excel function. The numerical amounts reflect the application of the formula to the values in rows 3 through 7. For example the value in cell B13 is the sum of the (teller hours in column D multiplied by the Lambda values λ_js in column G). Hence the B13 cell value equals (20)(1) + (30)(0) + (40)(0) +(20)(0) + (10)(0) = 20.

The next set of commands use the Excel INDEX function and simply place the actual input and output values of the service unit being evaluated by DEA on the right-hand-side of the equation in rows 13, 14, and 15 of column D.

For the unit under evaluation (unit B1), we enter the following formulas into cells D13:D15.

Cell D13 = $F$12*INDEX(D3:E7,E11,1)
Cell D14 = $F$12*INDEX(D3:E7,E11,2)
Cell D15 = INDEX(B3:B7,E11,1)

You can verify the above equations by noting that the values in cells D13, D14, and D 15 applying the above equations reflect B1's inputs (20 hours and $300 supplies) and B1's outputs (100 transactions).

The function INDEX(array,row number,column number) returns the value in the specified row and column of the given array. Because cell E11 contains the current value of 1, the INDEX function in cell D13 returns the value in first row and first column of the Input Used array D3:E7 (or the value in cell D3, the Teller Hours input for unit B1). When the value in cell E11 changes from 1 to 7, the INDEX functions in cells D13:D15 returns the input and output values for a specific service unit under evaluation. This feature becomes more

[4] Note that for Excel cells, these formulas are input with the first entry being the equal (=) sign. For example, in cell B13, enter the following =SUMPRODUCE(B3: .....).

obvious and useful when we provide the Visual Basic for Applications (VBA) code to automate the DEA computation.

To complete the setup of the DEA equations, we will use the Excel Solver and add the constraints corresponding directly to the equations (a) and (b) in model (2.2). In addition, we need to instruct the Solver about which cell to solve for and whether it should seek to *maximize* or minimize that cell. The instruction will initially be to *minimize* the cell F12, consistent with the dual equation model (2.2). Once these constraints and this command is incorporated into the program, a solve command in Excel will locate the DEA efficiency rating and the related λ values found in the bank branch example.

Note that we are using the dual linear program DEA model here and the instruction is to minimize the efficiency rating. If you are unfamiliar with linear programming, the term, "dual" and the concept of minimizing the efficiency rating may be confusing. If you are familiar with this, ignore the following explanation.

DEA does try to make every unit as efficient as possible, which is why we describe it as giving the benefit of the doubt to each service unit. This is a key strength of DEA from a management perspective as it allows only truly inefficient units to be identified as such. The initial DEA model (Table 2-5) does seek to maximize the efficiency value θ to make every unit look as good as possible with efficiency rating as close to 100% as possible in comparison with the other service units being evaluated. We use the "dual" program to solve (equation 2.1) the linear program because of technical computation advantages. However, the minimum θ value with the "dual" will equal the maximum θ value with the "primal". The mathematical explanation for this is basic to linear programming but beyond the scope of this volume.

2.4.3.1 Using Solver

After the DEA model is set up in the spreadsheet, we can use Solver to find the optimal solutions. First, we need to invoke Solver in Excel by using the Tools/Solver menu item, as shown in Figure 2-3.

If Solver does not exist in the Tools menu, you need to select Tools/Add-Ins, and check the Solver box, as shown in Figure 2-4. (If Solver does not show in the Add-Ins, you need to install the Solver first.)

Now, you should see the Solver Parameters dialog box shown in Figure 2-5.

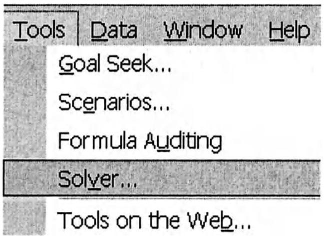


Figure 2-3. Display Solver Parameters Dialog Box

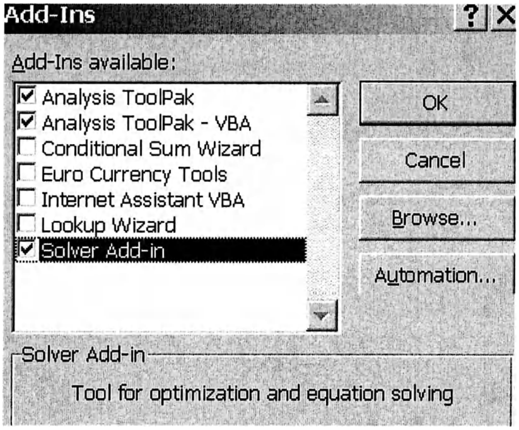


Figure 2-4. Solver Add-In

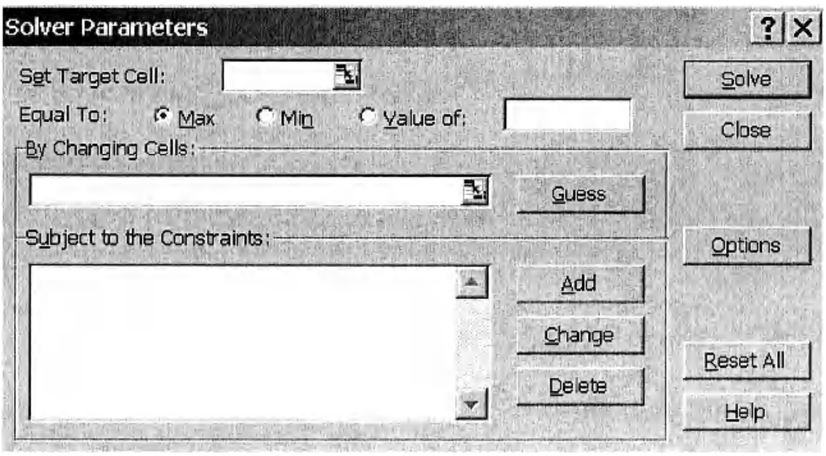


Figure 2-5. Solver Parameters Dialog Box

2.4.3.2 Specifying the Target Cell

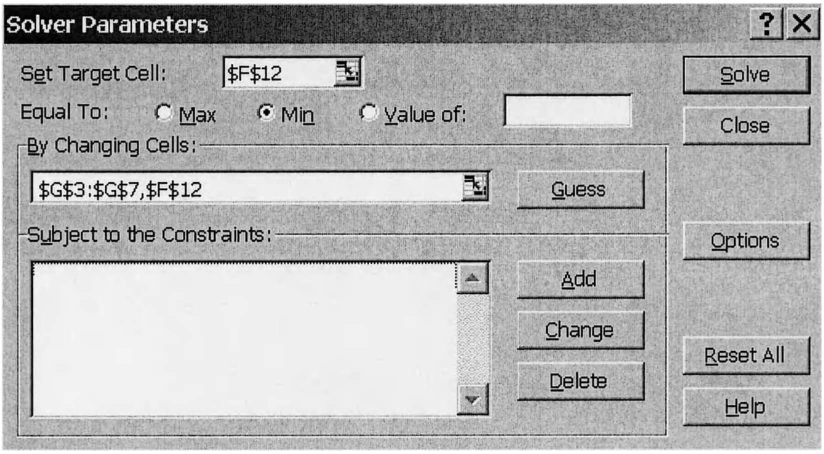


Figure 2-6. Specifying Target Cell and Changing Cells

Set Target Cell indicates the objective function cell in the spreadsheet, and whether its value should be maximized or minimized. In our case, the target cell is the DEA efficiency represented by cell F12, and its value should be minimized as indicated in model (2.2).

2.4.3.3 Specifying Changing Cells

Changing Cells represent the decision variables in the spreadsheet. In our case, they represent the λ_j (j = 1,2, , 7) and θ, and should be cells G3:G7 and F12, respectively (see Figure 2-6).

2.4.3.4 Adding Constraints

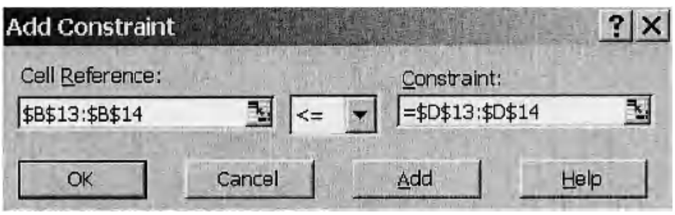


Figure 2-7. Adding Constraints

Constraints represent the constraints in the spreadsheet. In our case, they are determined by cells B13:B15 and D13:D15. For example, click the Add button shown in Figure 2-6, you will see the Add Constraint dialog box shown in Figure 2-7.

In the spreadsheet model shown in Figure 2-2, we have three constraints. The "Cell Reference" corresponds to the DEA Efficiency Reference Set, and "Constraint" corresponds to the service unit under evaluation. The first two constraints are related to the two inputs (see Figure 2-2). Click the Add button to add additional constraints (in this case, we have one output constraint), and click the OK button when you have finished adding the constraints. The set of the constraints are shown in Figure 2-9.

2.4.3.5 Non-Negativity and Linear Model

Note that λ_j and θ are all non-negative, and the model (2.2) is a linear programming problem. This can be achieved by clicking the Option button in Figure 2-6, and then checking the Assume Non-Negative and Assume Linear Model boxes, as shown in Figure 2-8. This action should be performed for each DEA model. In the rest of the book, we will not show the Solver Options dialog box, but please *be sure that the settings are consistent with Figure 2-8 each time you begin a new series of DEA evaluations.*

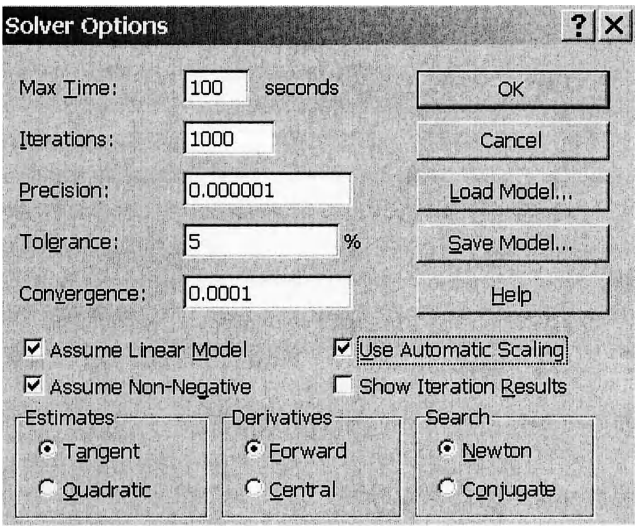


Figure 2-8. Non-Negative and Linear Model

When the Assuming Linear Model option is checked, Solver conducts a number of internal tests to see if the model is truly linear. When the data are poorly scaled, Solver may show that the conditions for linearity are not satisfied. To circumvent this, we may check the box of "Use Automatic Scaling" in the Solver Options dialog box.

2.4.3.6 Solving the Envelopment Model

Now, we have successfully set up the Solver Parameters dialog box, as shown in Figure 2-9. Click the Solve button to solve the model. When Solver finds an optimal solution, it displays the Solver Results dialog box, as shown in Figure 2-10.

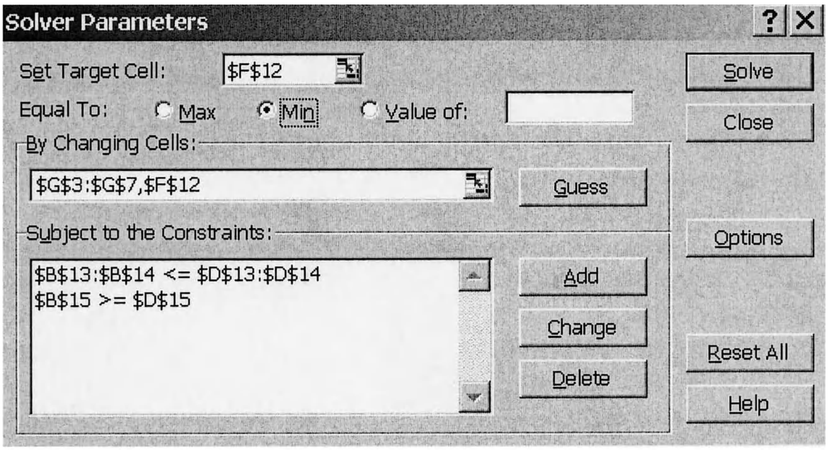


Figure 2-9. Solver Parameters for DEA Model (2.2)

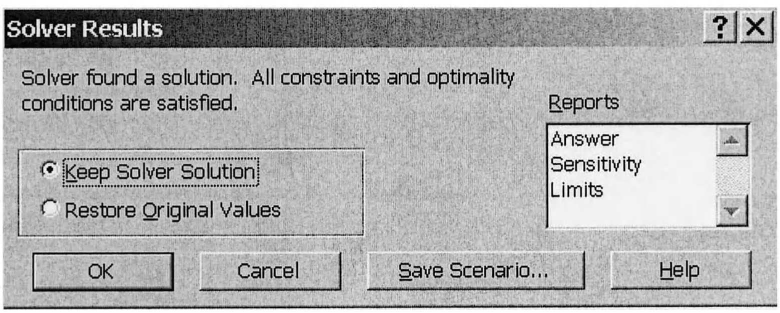


Figure 2-10. Solver Results Dialog Box

	A	B	C	D	E	F	G
1		Service Output		Input Used			
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)		λ
3	B1	1,000		20	300		0
4	B2	1,000		30	200		0
5	B3	1,000		40	100		0
6	B4	1,000		20	200		0.714286
7	B5	1,000		10	400		0.285714
8							
9							
10							
11		Efficiency Reference		unit under	1	Efficiency	
12	Constraints	Set		Evaluation		0.857143	
13	Teller Hours (H)	17.14285714	≤	17.14285714			
14	Supply Dollars (S)	257.1428571	≤	257.1428571			
15	Service Output	1000	≥	1000			

Figure 2-11. Results for Branch B1

Figure 2-11 shows the results for branch B1. Cell F12 indicates that the efficiency for branch B1 is 0.857. The optimal λ in cells G3:G7 indicates that B4 and B5 are in the efficiency reference set. The $\lambda_{B4}^{*} = 0.7143$ in cell G6 and $\lambda_{B5}^{*} = 0.2857$ in cell G7 indicate that if the input used by branch B1 is distributed among B4 and B5, then the efficiency can be improved. Note that these calculations are identical to the calculations in table 2-4.

2.4.3.7 Automating the DEA Calculation

To complete the analysis for the remaining 4 companies, one needs to manually change the value in cell E11 to 2, 3, 4, 5 and use Solver to re-optimize the spreadsheet model for each service unit and record the efficiency scores (in column H, for instance). When the number of service units becomes large, the manual process is cumbersome.

Note that exactly the same Solver settings will be used to find the optimal solutions for the remaining service units (or DMUs). This allows us to write a simple Visual Basic VBA code to carry out the process automatically.

Before we write the VBA code, we need to set a reference to Solver Add-In in Visual Basic (VB) Editor. Otherwise, VBA will not recognize the Solver functions and you will get a "Sub or function not defined" error message.

We may follow the following procedure to set the reference. Enter the VB Editor by pressing *Alt-F11* key combination (or using the Tools/Macro/Visual Basic Editor menu item). Open the Tools/References menu in the VB Editor. This brings up a list of references. One of these should be **Solver.xla** (see Figure 2-12).

To add the reference, simply check its box. If it says "Missing: Solver.xla", then click the Browse button and search for **Solver.xla**. If you are using Excel XP, the **Solver.xla** is usually located at C:\ Program Files\ Microsoft Office\ Office10\ Library\ Solver. Otherwise, the **Solver.xla** is usually located at C:\ Program Files\ Microsoft Office\Office\Library\ Solver. However, this depends on where the Microsoft Office is installed.

After the Solver reference is added, we should see "Reference to Solver.xla" under the "References" in the VBA Project Explorer window shown in Figure 2-13. (The file "Table2-2 spreadsheet.xls" in the CD contains the spreadsheet model.)

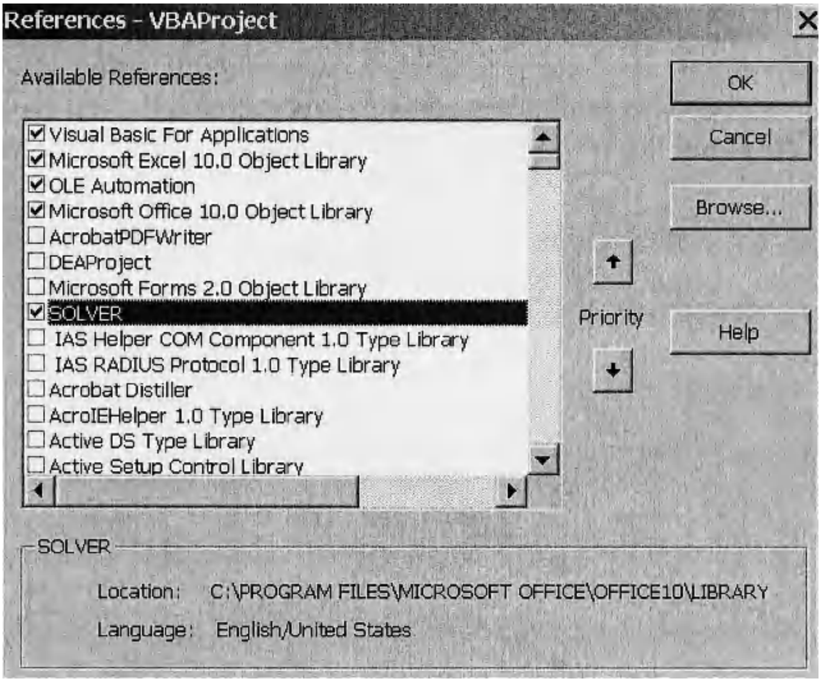


Figure 2-12. Adding Reference to Solver Add-In

Next, select the Insert/Module menu item in the VB Editor (Figure 2-12). This action will add a Module (e.g., Module1) into the Excel file. (You can change the name of the inserted module in the Name property of the module.)

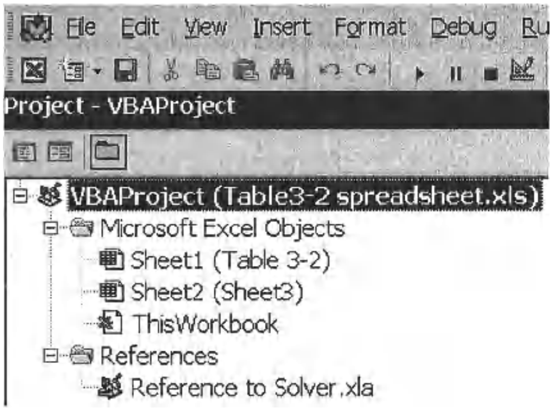


Figure 2-13. Reference to Solver Add-In in VBA Project

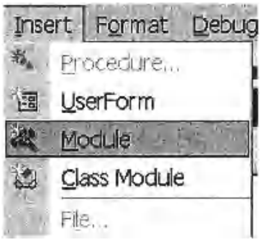


Figure 2-14. Insert a Module

Now, we can insert the VBA code into the Module1. Type "Sub DEA ()" in the code window (see Figure 2-15). This generates a VBA procedure called DEA, which is also the Macro name (see Figure 2-16). Figure 2-15 shows the VBA code for automating the DEA calculation.

To insert this code, one could just type in the commands in Figure 2-15. An electronic version of this set of commands is available in the "Table2-2 spreadsheet.xls" in attached CD.

The Macro statement "SolverSolve UserFinish:=True" tells the Solver to solve the DEA problem without displaying the Solver Results dialog box. The "Offset(*rowOffset*, *columnOffset*)" property takes two arguments that correspond to the relative position from the upper-left cell of the specified Range. When we evaluate the first service unit, i.e., DMUNo = 1, Range("H2").Offset(1,0) refers to cell H3. The statements "With Range("H2") and ".Offset(DMUNo, 0) = Range("F12") take the optimal objective function value (efficiency score) in cell F12 and place it in cell H "DMUNo+2" (that is, cell H3, H4, ..., H7).

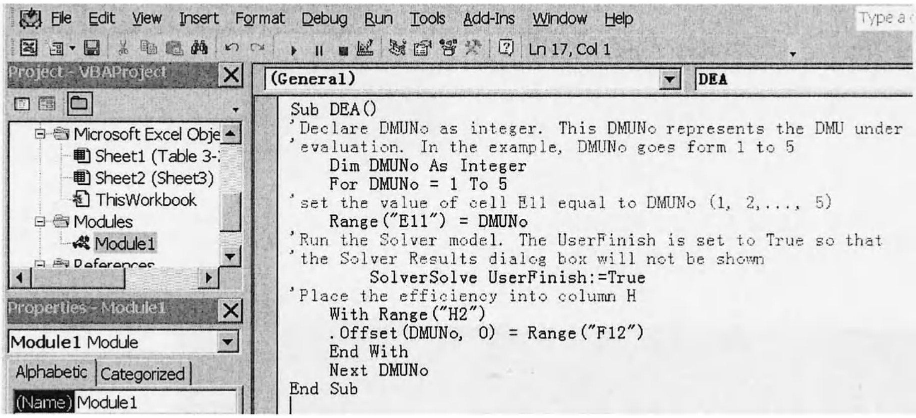


Figure 2-15. VBA Code for Envelopment Model

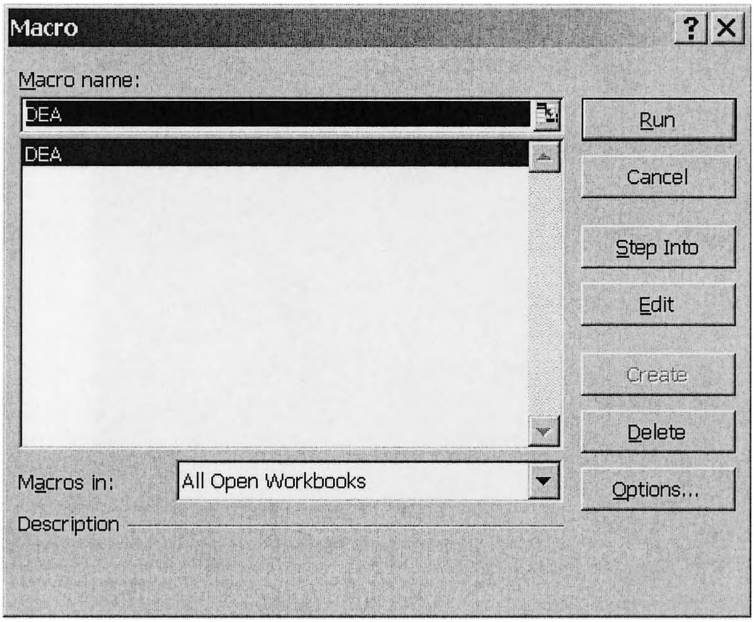


Figure 2-16. Run "DEA" Macro

	A	B	C	D	E	F	G	H
1		Service Output		Input Used				
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars		λ	Efficiency
3	B1	1,000		20	300		0	0.857143
4	B2	1,000		30	200		0	0.857143
5	B3	1,000		40	100		0	1
6	B4	1,000		20	200		0	1
7	B5	1,000		10	400		1	1
8								
9								
10								
11		Efficiency Reference		unit under	5	Efficiency		
12	**Constraints**	Set		Evaluation		1		
13	Teller Hours (H)	10	≤	10				
14	Supply Dollars (S)	400	≤	400				
15	Service Output	1000	≥	1000				

Figure 2-17. Efficiency Scores for the Five Bank Branches

Enter the Run Macro dialog box by pressing *Alt-F8* key combination (or using the Tools/Macro/Macros menu item). You should see "DEA", as shown in Figure 2-16. Select "DEA" and then click the Run button. This action will generate the efficiency scores (cells H3:H7) for the 5 bank branches, as shown in Figure 2-17.

Note that the efficiency scores in column H of Figure 2-17 are consistent with the observed inefficiencies in Figure 2-1 and the DEA results reported in Table 2-3. the Excel solver has evaluated each service unit. The last service unit evaluated was the efficient unit, B5, which the reason the efficiency rating in cell F12 is 1.0, signifying that B5 is efficient and this is mirrored in cell H7.

The previous macro "DEA" does not record the optimal λ_j values in the worksheet. This can be done by the adding a VBA procedure named "DEA1" into the existing module.

```
Sub DEA1()
'Declare DMUNo as integer. This DMUNo represents the DMU under
'evaluation. In the example, DMUNo goes form 1 to 5

    Dim DMUNo As Integer
    For DMUNo = 1 To 5
'set the value of cell E11 equal to DMUNo (1, 2,..., 5)
    Range("E11") = DMUNo

'Run the Solver model. The UserFinish is set to True so that
'the Solver Results dialog box will not be shown
    SolverSolve UserFinish:=True

'Place the efficiency into column H

    Range("H" & DMUNo + 2) = Range("F12")
```

```
'Select the cells containing the optimal lambdas

    Range("G3:G7").Select

'copy the selected lambdas and paste them to row "DMUNo+2"
'(that is row 3, 4, ..., 7) starting with column J

    Selection.Copy
    Range("J" & DMUNo + 2).Select
    Selection.PasteSpecial Paste:=xlPasteValues,
Transpose:=True

    Next DMUNo
End Sub
```

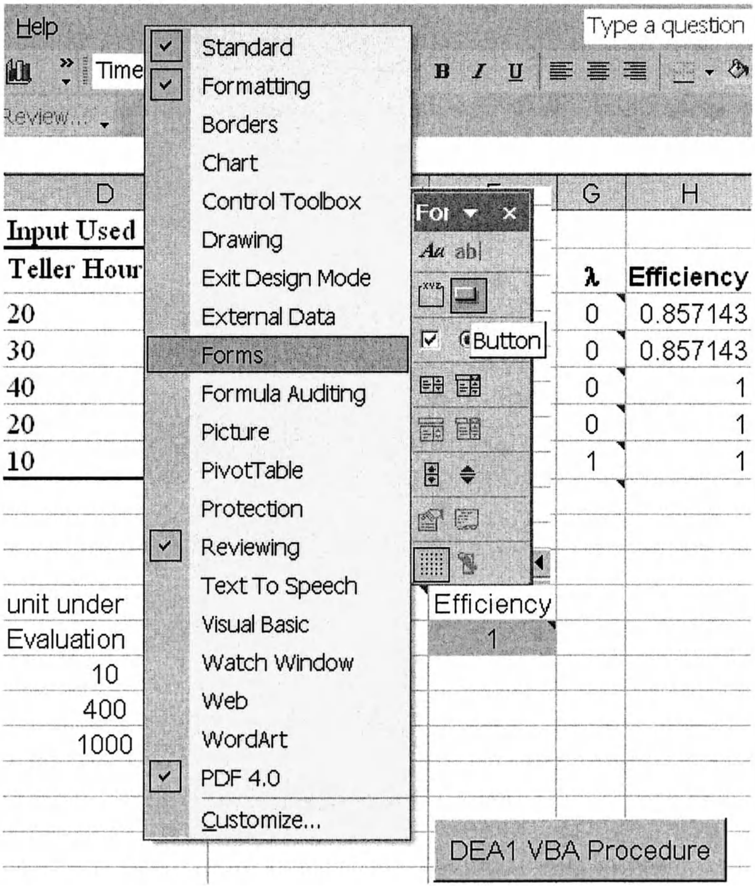


Figure 2-18. Adding a Button with Macro

In the Run Macro dialog box, select "DEA1" and then click the Run button. The procedure "DEA1" will record both the efficiency scores and the related optimal values on λ_j (j = 1, 2, , 5) (see file "Table2-2 spreadsheet.xls" in the CD). This will generate the λ_j values for ERS units such as the values reported in Figure 2-11 where the values for unit B1's ERS branches are reported as 0.713286 for ERS unit B4 and 0.285714 for ERS unit B5. In the solution worksheet, λ_j values will appear in row 3 starting column J for the first service unit.

We can also create a button to run the VBA procedure. First, we select the View/Toolbars/Forms menu item (or right-click on any toolbar in Excel) (see Figure 2-18). The fourth item on the Forms toolbar is for creating buttons to run macros (VBA procedures). Click and drag the button onto your worksheet containing the DEA spreadsheet and the Solver parameters. You will immediately be asked to assign a macro to this button. Select "DEA1". At this point, the button is selected. To run the selected macro, you have to deselect the button by clicking anywhere else on the worksheet. You can always assign a different macro to the button by right clicking on the button and selecting "Assign Macro".

2.4.4 Solving Multiplier DEA Model as a Linear Program in Spreadsheets

We now demonstrate how the multiplier model (2.1) can be solved via Solver. This is the original configuration of DEA in contrast to the dual model (2.2) used above. The difference is the dual solves for the λ lambda values while the multiplier model solves for the u and v (use the real variable notations) weights that are applied to the inputs and outputs. Figure 2-19 presents the multiplier spreadsheet model. Cells B11 and C11:D11 are reserved for the output and input multipliers (weights). They are the changing cells in the Solver parameters. Cell B12 is reserved to represent the service unit under evaluation.

The target cell is B13, which represents the efficiency – weighted output for the DMU under evaluation. Its formula is "=$B$11*INDEX(B3:B7,B12,1)".

	A	B	C	D	E	F
1		Service Output		Input Used		
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)	Constraints
3	B1	1,000		20	300	-0.142857143
4	B2	1,000		30	200	-0.285714286
5	B3	1,000		40	100	-0.428571429
6	B4	1,000		20	200	0
7	B5	1,000		10	400	0
8					weighted input	1
9						
10						
11	Multipliers	0.000857143		0.028571429	0.001428571	
12	unit under evaluation	1				
13	Efficiency	0.857142857				

Figure 2-19. Multiplier Spreadsheet Model

Cell F3 contains the formula "=$B$11*B3 – SUMPRODUCT ($D$11:$E$11,D3:E3)" which represents the difference between weighted output and weighted input for branch B1. The formula in cell F3 is copied into cells F4:F7. These values will be set as non-positive in the Constraints of Solver parameters (see Figure 2-20).

The formula for cell F8 is "=SUMPRODUCT ($D$11:$E$11, INDEX(D3:E7,B12,0))". The value of cell F8 will be set equal to one in the Constraints of Solver parameters (see Figure 2-20).

The function INDEX(array,row number,0) returns the entire row in the array. For example, the value for cell B12 is one, therefore INDEX(D3:E7,B12,0) returns the first inputs across all units, i.e., cells D3:E3.

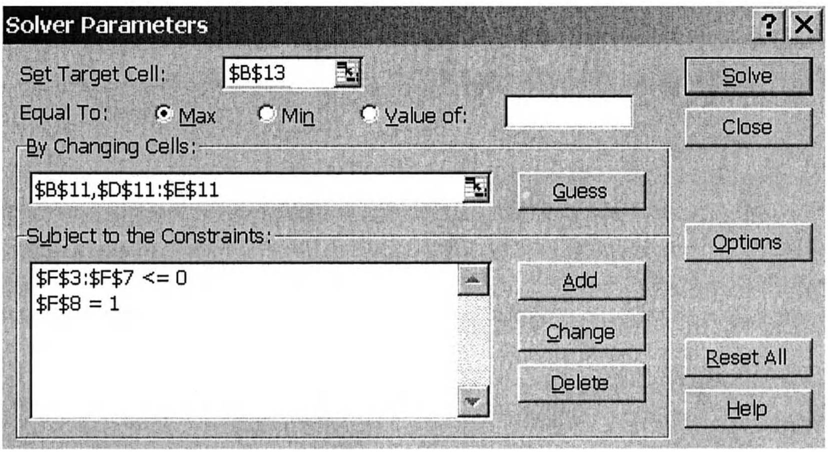


Figure 2-20. Solver Parameters for Multiplier Model

	A	B	C	D	E	F
1		Service Output		Input Used		
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)	Constraints
3	B1	1,000		20	300	-0.28571429
4	B2	1,000		30	200	-0.14285714
5	B3	1,000		40	100	0
6	B4	1,000		20	200	0
7	B5	1,000		10	400	-0.42857143
8					weighted input	1
9						
10						
11	Multipliers	0.000857143		0.014285714	0.002857143	
12	unit under evaluation	2				
13	Efficiency	0.857142857				

Figure 2-21. Efficiency for Branch B2

Figure 2-20 shows the Solver parameters for the spreadsheet model in Figure 2-19. The solution in Figure 2-19 is actual optimal for branch B1. To calculate, we change the value of cell B12 to 2 and re-solve the problem. Figure 2-21 shows the results for branch B2. Note that we have u = 0.0857 ($\times 10^{-2}$) in cell B11, v1 = 1.429 ($\times 10^{-2}$) in cell D11 and v2 = 0.286 ($\times 10^{-2}$) in cell E11.

We can also insert a VBA procedure "Multiplier" to automate the computation.

```
Sub Multiplier()
    Dim i As Integer
    For i = 1 To 5
        Range("B12") = i
    SolverSolve UserFinish:=True
'record the efficiency scores
    Range("G1").Offset(i + 1, 0) = Range("B13")
'record the optimal multipliers
    Range("B11").Copy
    Range("H1").Offset(i + 1, 0).Select
    Selection.PasteSpecial Paste:=xlPasteValues
    Range("D11:E11").Copy
    Range("I1").Offset(i + 1, 0).Select
    Selection.PasteSpecial Paste:=xlPasteValues
    Next i
End Sub
```

This VBA procedure takes the efficiency in cell B13 and places it into cells G3:G7, and also takes the optimal multipliers and places them into cells H3:H7 for the output and I7:J7 for the inputs for the 5 units. Select and run the macro "Multiplier" in the Run Macro dialog box will generate the efficiency results. You may also create a button in Forms toolbar and assign macro "Multiplier" to the button (see file "Table2-2 spreadsheet.xls" in the CD).

The results from model (2.2) are reported in Figure 2-22. These multipliers (or weights) correspond to the weights in the DEA model shown in Table 2-5.

	A	G	H	I	J
1					
2	Service Unit	Efficiency	T	H	S
3	B1	0.857143	0.000857	0.02857	0.00143
4	B2	0.857143	0.000857	0.01429	0.00286
5	B3	1	0.001	0	0.01
6	B4	1	0.001	0.01667	0.00333
7	B5	1	0.001	0.03333	0.00167

Figure 2-22. Optimal Multipliers

2.5. CONCLUSIONS

We have described the basic DEA model, the insights it provides in a simple example where similar insight could be gained from observation and where the linear programming calculations could easily be done without a computer. The following chapter illustrates the power of DEA with multiple inputs and outputs. In addition, chapter 5 describes a variety of models to adapt DEA for particular applications and circumstances to enable users to employ the most appropriate models for their organization. Chapter 4 introduces a ready-to-use Excel Solver-based DEA software called *DEAFrontier*. The *DEAFrontier* does not set limit on the number of units, inputs or outputs. With the capacity of Excel Solver, the *DEAFrontier* can deal with large sized performance evaluation tasks.[5]

[5] The capacity of a DEA to handle large numbers of inputs, outputs and service units is only constrained by the computing power available to the user. After specifying the model and after assembling and cleaning the data, DEA has been applied to organizations with over 1500 service units without any more computational effort that was required for the examples in this chapter. The results are naturally more voluminous, but they can also generate substantial benefits. The bank example noted at the start of this chapter identified well over $100 million of excess resources based on a DEA analysis of over 1500 branch units.

Chapter 3

DEA CONCEPTS FOR MANAGERS

Applying and managing productivity with DEA

3.1. INTRODUCTION

Chapter 2 presents the basic DEA model and describes how DEA can be run using Microsoft® Excel, and suggests how accessible DEA evaluations are. The examples used in the prior chapter were purposely uncomplicated and simple enough to allow one to analyze the service units and conclude about their performance without DEA. This chapter increases the complexity of the examples to illustrate the types of insights available from DEA as the data approach applications to real sets of service provider. In addition, the interpretation of the data from a management perspective including the unique insights provided to improve performance and limitations of these insights is emphasized to help readers evaluate the usefulness of DEA for their organization.

The following section begins by developing the more complex example and introducing additional DEA characteristics. Specifically, DEA continues to identify efficient best practice units but differentiates between strong efficiency units that are true best practice models for other service units and a group of weakly efficient units that are not inefficient compared to other units but are also not necessarily best practice models for other service units.

3.2. DEA EFFICIENT AND WEAKLY EFFICIENT: CONCEPTS AND EXAMPLES

We begin with the same five branches (B1 through B5) and add seven other branches. This use of DEA might be employed for example when the bank has acquired another bank with seven branches (B6 through B12). Table 3-1 lists the new set of output and input data for these branches and the DEA results. These branches are also plotted on a two-dimension graph (Figure 3-1).

Table 3-1. Extended Bank Branch Example

Branch	Output Transactions	Input 1 Hours (H)	Input 2 Supplies (S)
B1	1000	20	300
B2	1000	30	200
B3	1000	40	100
B4	1000	20	200
B5	1000	10	400
B6	1000	15	350
B7	1000	31	250
B8	1000	50	100
B9	1000	45	100
B10	2000	80	200
B11	2000	18	750
B12	2000	45	410

The new branches have characteristics that are intentionally different from those of the initial five presented in Table 2-3 in chapter 2. Branches 6 and 7 are not totally dominated by any other branch as were branches 1 and 2. That is, B1 used the same input hours as B4 but used more supplies; B1 was thus absolutely less efficient than B4, which is referred to as dominated. In contrast, B6 has less H and more S than B4 but it also uses more H but less S than B5 (and B11). Hence, it is not purely dominated by B4, B5 (or B11). (B7 uses one more hour and 50 more supply dollars than B2 and therefore is dominated by B2, which is an inefficient unit itself).

Branches 8 and 9 are units that use the same amount of one input and more of the other input and which define a new "facet" in the best practice frontier. These units are clearly not efficient compared to one other branch, in this case B3.

Branches 10, 11, and 12 have 2,000 transactions, twice the output of the other branches. (In Figure 3-1, the inputs of these branches are divided by 2.) Branch 10 has exactly two times the inputs of B2 so it is equally efficient with no scale economies present. B11 has scale economies in that its inputs are less than twice those of B5. B12 has scale diseconomies in that its inputs are more than twice those of B4.

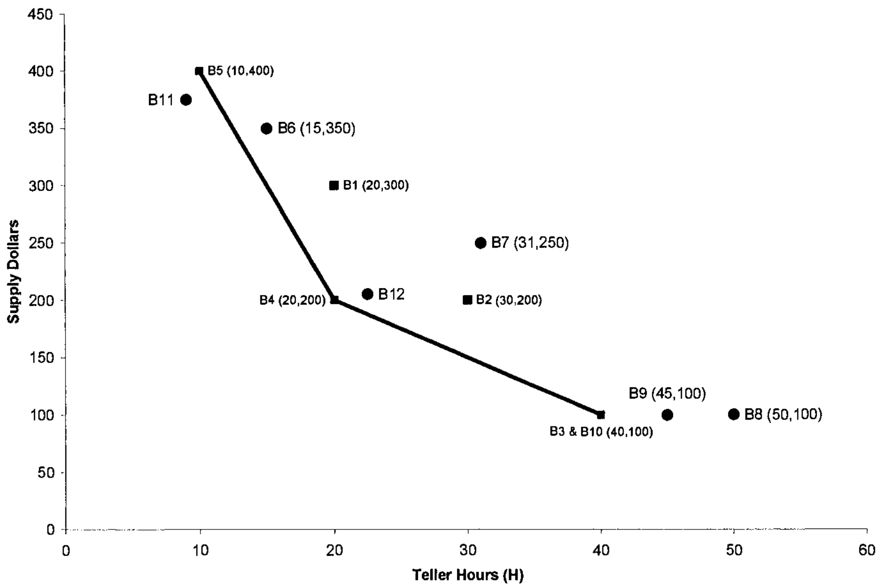


Figure 3-1. Extended Bank Branch Example

Note that when DEA compares B11 with B5, B5 is less efficient than B11 ÷ 2 (see Table 3-1). Consequently, B11 becomes the best practice standard and not B5.

Note that with only twelve branches, two inputs and one output, it is already difficult to identify the inefficiencies present by pure observation. Ratio analysis would not be sufficient to evaluate the branch performance without applying some weights for the relative cost or value of supplies and hours.

Applying DEA model (2.2) in chapter 2, we have the results shown in Table 3-2.

The file "Table 3-1 spreadsheet.xls" in the CD provides the spreadsheet model for the branches in Table 3-1 and the VBA (Visual Basic for Applications) procedure "DEA2". This VBA procedure "DEA2" is similar to the VBA procedure "DEA1" provided in chapter 2.

Table 3-2. DEA Results for the Extended Bank Branch Example

Branch	Original DEA rating	DEA efficiency	Efficiency reference set (λ)
B1	0.857	0.838	B4(0.706),B11(0.147)
B2	0.857	0.857	B3(0.286),B4(0.714)
B3	1	1	
B4	1	1	
B5	1	0.927	B4(0.024),B11(0.488)
B6		0.880	B4(0.382),B11(0.309)
B7		0.740	B3(0.148),B4(0.852)
B8		1	B3 (1)
B9		1	B3 (1)
B10		1	
B11		1	
B12		0.945	B3(0.126),B4(1.874)

Several observations suggest aspects of DEA capabilities and behavior not apparent from the prior example.

When the observation set of service units increases, these new units may be more or less efficient than the original units. While B1 and B2 continue to be rated as inefficient, B2 has the same rating while B1 has a lower rating of 0.831 vs. 0.857 in the original five-branch example. This is due to the introduction of B11, which was more efficient than B5. B1 is now compared to B4 and B11 and is more inefficient than it appeared when B11 was excluded. Note that when service units are eliminated from the set of service units being evaluated, the efficiency rating of an inefficient branch will stay the same or increase. By eliminating B11 from this set, B1 would appear more efficient because B11 is not present as a basis for comparison. B5 was originally efficient and now appears inefficient compared with B11 and B4. This is consistent with the way B11 was constructed. That is, B11 has scale economies. B5 is essentially being compared to half of B11's outputs and inputs.

Note that B8 and B9 comprise a new facet in the DEA frontier portrayed in Figure 3-1. They are both found to be inefficient compared with B3 as one would anticipate. Specifically, B8 and B9 should reduce their Teller Hours input by 10 and 5 units, respectively. These individual reductions are called "DEA slacks" and the calculation of these with DEA for B8 will be illustrated below.

Mathematically, to obtain these "DEA slacks", one needs to solve the following linear programming problem after model (2.2) in chapter 2 is solved.

$$\max \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$$
subject to
$$\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = \theta^* x_{io} \qquad i = 1,2,\ldots,m; \tag{3.1}$$
$$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{ro} \qquad r = 1,2,\ldots,s;$$
$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n.$$

where θ^* is the DEA efficiency score obtained from model (2.2) in chapter 2, and s_i^- and s_r^+ are input and output slacks, respectively.

Applying model (3.1) to B8 yields,

$$\max s_H^- + s_S^- + s_T^+$$
subject to
$$20\lambda_1 + \quad 30\lambda_2 + \ldots + \quad 50\lambda_8 + \ldots + \quad 45\lambda_{12} + s_H^- = 1 \bullet 50$$
$$300\lambda_1 + \quad 200\lambda_2 + \ldots + \quad 100\lambda_8 + \ldots + \quad 410\lambda_{12} + s_S^- = 1 \bullet 100$$
$$1000\lambda_1 + 1000\lambda_2 + \ldots + 1000\lambda_8 + \ldots + 2000\lambda_{12} - s_T^+ = 1000$$
$$\lambda_j \geq 0 \;\; (j = 1,\ldots,12)$$

which yields an optimal solution of $s_H^{-*} = 10$ and $s_S^{-*} = s_T^{+*} = 0$.

Note that for B8, the efficiency reference set includes only B3 and the λ for B3 = 1 and λ value for the other 11 branches is 0. (Only efficiency reference set service units will have non-zero λ values). Hence the equations for the above calculation are

$$\max s_H^- + s_S^- + s_T^+$$
subject to
$$40\lambda_3 + s_H^- = 1 \bullet 50 = 50$$
$$100\lambda_3 + s_S^- = 1 \bullet 100 = 100$$
$$1000\lambda_3 - s_T^+ = 1000$$
$$\lambda_j \geq 0 \;\; (j = 1,\ldots,12)$$

It is obvious that setting λ for B3 = 1 in the above equations that the slack for teller hours (s_H^-) is 10, the amount of excess teller hours used by B8. Similarly the slack for supplies (s_s^-) and transactions (s_T^+)

is zero. This slack calculation is similar to the excess resources calculated in Chapter 2 for Branch 2 in Table 2-3.

Thus, a complete DEA calculation for the envelopment model (e.g., model (2.2) in chapter 2) involves two stages: first, calculate the efficiency score, θ^*, then calculate the slacks while fixing the θ^*. (Both stages are calculated automatically by some DEA codes, so that sequential calculations appear to be simultaneous. The *DEAFrontier* software provided in this book calculates the two stages separately.)

Definition 3.1 (DEA Efficient): The performance of DMU_o is fully (100%) efficient if and only if both (i) $\theta^* = 1$ and (ii) all slacks $s_i^{-*} = s_r^{+*} = 0$.

Definition 3.2 (Weakly DEA Efficient): The performance of DMU_o is weakly efficient if and only if both (i) $\theta^* = 1$ and (ii) $s_i^{-*} \neq 0$ and/or $s_r^{+*} \neq 0$ for some i and r.

For the example in Table 3-1, B3, B4, B10 and B11 are DEA efficient and B8 and B9 are weakly DEA efficient. All the remaining branches are inefficient.

From a managerial perspective, the classification of efficient service units as DEA efficient and weakly DEA efficient has various implications. (In practice, other terms might be used to characterize these categories, as the term weakly efficient may be misinterpreted if the user is not acquainted with the technical definition of the term.)

1. The weakly efficient units will not be in the efficient reference set of other service units if we adopt the two-stage process of calculating models (2.2) and (3.1). If we only run model (2.2), then weakly efficient units can still be in the efficient reference set of other service units. This is because their efficiency score is equal to one. In any case, the weakly efficient units are not themselves influencing the efficiency ratings of other service units. If they are removed from the data set, the other analytical information from DEA for all the other service units will be unchanged.
2. The process of analyzing the DEA results requires considering whether efficient reference set service units are accurately specified. This is critical because these efficient reference set units define the degree of inefficiency in other units. In contrast, "weakly" efficient units do not influence the efficiency ratings of other units and it is not necessary to

scrutinize these before relying on the DEA efficiency evaluation of the other service units.

3. These units define an extreme facet by themselves and may warrant independent analysis of their performance. DEA provides very explicit information about the units that are "DEA efficient" and the units that are inefficient. The least amount of insight from DEA is provided on these "weakly" efficient units.

For inefficient service units, the following formula provides a path to the efficient frontier

$$\begin{cases} \hat{x}_{io} = \theta^* x_{io} - s_i^{-*} & i = 1,2,\ldots,m \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r = 1,2,\ldots,s \end{cases} \tag{3.2}$$

where θ^* is the efficient score obtained from model (2.2) and s_i^{-*} and s_r^{+*} are obtained from model (3.1). Formula (3.2) provides an efficient target for a specific inefficient service unit.

For example, for B9, the efficiency score is 1, and s_H^{-*} = 5. Thus, applying (3.2) to B9 yields B3 as the efficient target.

Now, if we consider the dual program to model (2.2), i.e., the multiplier model (2.1), the nonzero slacks in model (3.1) indicate that B8 and B9 assign a zero value to the input they over utilized - in this case, hours (H). i.e., B8 and B9 would appear to be as efficient as B3, because the DEA efficiency score is equal to one. This would be tantamount to saying that input H had no value, was cost free, or was unimportant. For this reason, DEA includes a constraint to make these weights non-zero. Similarly, outputs also are not allowed to have a zero weight, as that would suggest that the output was valueless, unimportant, and/or irrelevant.

The excess resources used by B8 and B9 (DEA weakly efficient) can be directly read from the model (3.1). This indicates that B8 used ten excess units of H and B9 used five excess units of H, as is apparent from Table 3-1 and the graph in Figure 3-1.

Scale economies in basic DEA analysis: The three branches with 2,000 output units, B10, B11, and B12, are rated consistent with their designed behavior intended in the design of this hypothetical dataset to illustrate the impact of scale economies using the basic DEA formulation. B10 is as efficient as B3 and both are rated 100%

efficient – B10 has exactly twice the output and uses twice the inputs of B3. B11 is more efficient than B5 and is rated as efficient. B12 is less efficient than B4 and is rated as 94.5% efficient compared to B3 and B4, reflecting inefficiency, which in this case relates to diseconomies of scale.

When only scale economies exist, the large more efficient units appear inefficient as is the case for B11 vs. B5. (B5 produces half the service output of B11 but uses more than half the amount T and H resources used by B11.) Where no scale economies exist, larger and smaller units will appear to be efficient with DEA unless other inefficiencies are also present. Where scale diseconomies are present, the smaller units will appear as more efficient than larger units, as was the case for B12.

Note that when scale and other inefficiencies are present, additional analysis is required to separate scale and technical inefficiencies. One simple approach is to separate and compare similar size units. Other approaches to interpret the scale economies will be discussed at the end of this chapter. The more sophisticated approaches will be meaningful only if there are many service units of all sizes in the data set.

In this simple example, no conclusion can be reached about whether large units should or should not be more efficient than small units. Any such assertions would be unfounded and possibly misleading. For example, DEA reports B3 and B10 to have equal efficient ratings of 1.0 or 100%. If the type of service units being evaluated were known to be able to benefit of economies of scale, then this would suggest that the larger unit B10 should be able to use fewer inputs per unit of service than small units and it may be operating less efficiently than B3. The only conclusion at this stage that are supportable with the analysis at this stage are that B5 is inefficient compared with B4 and large branch B11, and that B12 is inefficient compared with small branches B5 and B4. (In short, inefficiencies in the units of different sizes cannot be assumed to reflects scale economies. Inefficiencies may also be due to poor management and/or environmental factors unrelated to size or volume of activity.)

We finally note that the interpretation of the DEA output data is the same as with the example in Chapter 2. Two examples will be presented below to further familiarize the reader with these interpretation issues.

Branch 1 is now inefficient compared to B4 and B11 (see Table 3-2 for DEA output). One route for B1 to become efficient is to reduce hours by 3.2 and supplies by $48.53. (To calculate these amounts, the same calculation illustrated in Chapter 2, Table 2-4 would be completed to compare the composite of B4 and B11 to inefficient unit B1. For example, using the B1 DEA output, multiply the hour (H) input of B4 by its λ and B11 by its λ. That is (20 hours) X (0.706) =14.1 for B4 and (18 hours) X (0.147) = 2.7 for B11. Summing these weighted hours for B4 and B11, the total is 16.8, which is about 3.2 less than the hours used by B1.)

	A	B	C	D	E	F	G	H	I	J
1		Service Output		Input Used						
2	Service Unit	Transactions (T)		Hours (H)	Supply (S)	Constraints	Efficiency	T	H	S
3	B1	1,000		20	300	-0.1574803	0.838235	0.000838	0.02574	0.00162
4	B2	1,000		30	200	-0.0787402	0.857143	0.000857	0.01429	0.00286
5	B3	1,000		40	100	0	1	0.001	0	0.01
6	B4	1,000		20	200	0	1	0.001	0.01667	0.00333
7	B5	1,000		10	400	-0.2362205	0.926829	0.000927	0.02846	0.00179
8	**B6**	**1,000**		**15**	**350**	-0.1968504	0.880309	0.00088	0.02703	0.0017
9	**B7**	**1,000**		**31**	**250**	-0.1653543	0.740741	0.000741	0.01235	0.00247
10	**B8**	**1,000**		**50**	**100**	-0.0787402	1	0.001	0	0.01
11	**B9**	**1,000**		**45**	**100**	-0.0393701	1	0.001	0	0.01
12	**B10**	**2,000**		**80**	**200**	0	1	0.0005	0	0.005
13	**B11**	**2,000**		**18**	**750**	-0.3779528	1	0.0005	0.01535	0.00096
14	**B12**	**2,000**		**45**	**410**	-0.0551181	0.944882	0.000472	0.00787	0.00157
15					weighted input	1				
16										
17										
18	Multipliers	0.000472441		0.007874	0.001574803					
19	unit under evaluation	12								
20	Efficiency	0.94488189			Run Model (3.1)					
21										

Figure 3-2. Optimal DEA Multipliers

Other routes are also indicated by the DEA model (2.1). Figure 3-2 shows the spreadsheet model for model (3.1) with the data in Table 3-1 where a set of optimal multiplier values are reported in cells H3:J14. (See file "Table 3-1 spreadsheet.xls" for details.)

For B1, we have, u_T (Transactions) = 0.00083, v_h (Hours) = 0.0257, and v_s (Supply $'s) = 0.0016

This indicates the impact that one unit change in either input or output will have on the efficiency rating. B1's efficiency ratio is 0.838 or 83.8%. It must increase its efficiency rating by 100-83.8% or 16.2% to become relatively efficient. This can be translated into the following choices:

1. Increase transaction outputs by (0.162)/.00083 = 195 transactions,
2. Decrease input hours by (0.162)/.0257 = 6.3 hours,

3. Decrease input supplies by (0.162)/.00162 = \$100.00 which is the most apparent adjustment since B4 uses \$100 less supplies and the same input,
4. Increase and/or decrease the S and H inputs and output such that the net adjustment to the efficiency rating = 0.162.

Table 3-3. Potential Resource Saving and Inefficiency in Branch B12 Calculated by DEA

Outputs	Outputs & Inputs of B3	Outputs & Inputs of B4	Efficiency Reference Set for Service Unit B12

$$
\begin{array}{l}
\text{Transaction Processed(T)} \\
\underline{\text{Inputs}} \\
\text{Teller Hours (H)} \\
\text{Supply \$ (S)}
\end{array}
\quad (0.126)\times\begin{bmatrix}1{,}000\\ \\ 40\\ 100\end{bmatrix} + (1.874)\times\begin{bmatrix}1{,}000\\ \\ 20\\ 200\end{bmatrix} = \begin{bmatrix}2{,}000\\ \\ 42.52\\ 387.4\end{bmatrix}
$$

The composite for B2 can then be compared with the inefficient unit B2 as follows:

	Column 1 Composite Outputs & Inputs (from above)	**Column 2** Branch B12 Actual Outputs & Inputs	Column 2 – Column 1
(T)	1,000	2,000	0
(H)	42.52	45	2.48
(S)	387.4	410	22.60 $\left\{\begin{array}{l}\text{Excess Inputs}\\ \text{Used by}\\ \text{Branch B2}\end{array}\right.$

One caveat is that only a few of the efficiency improvement options are likely to be practical. While one can define a large set of mathematical choices, many of them will be meaningless or unfeasible from a managerial perspective. In practice, the managers responsible for the service units have selected the path they believe is most appropriate, which may not be the one that will lead to the greatest improvement in performance. Reasons for this are discussed in later chapters describing implementation of DEA in service organizations. (We should also note that the multipliers or weights are usually not unique.)

The second example will focus on branch 12 which is 94.5% efficient compared to its ERS of B4 and B3. Using (3.2), Table 3-3 reflects the composite of B4 and B3 compared with B12. It suggests that B12 should reduce its hours by 2.48 and supplies cost by $22.60 to become efficient.

Note that B12 can also become as efficient as B4 by reducing H by 5 hours and S by $10. This adjustment would make B12 as efficient as B4, since it would then have exactly twice the outputs and inputs of B4.

Further interpretation of the technical meaning of the DEA results will be discussed in the more complex two input-two output example which follows.

3.3. TWO INPUT—TWO OUTPUT EXAMPLE

The following example begins to replicate the complexity and analytical capabilities of DEA in a real application. DEA is applied to a data set where it is much more difficult to observe the inefficiencies. The data set of service units use two inputs and provide two distinct types of services. As you read this example, consider that DEA can handle more complex examples and provide similar insights. Managed care organizations may apply this to thousands of physicians treating hundreds of illnesses and using ten or more types of resources. Banks have branch networks of up to 8000 branches in India and up to 6000 in the US and use five to ten types of resources to provide 20 or more services. Even organizations with as few as 20 service units providing a variety of services using a variety of resources have found that the complexity required a technique such a DEA to detect areas where performance could be improved and pockets of poor performance.

The example also clarifies relative strengths of DEA over ratio analysis. While we do not compare other performance management techniques directly in this example, the reader is encouraged to consider the techniques they currently use and consider whether the DEA would complement those techniques, whether DEA could replace existing methods, or whether there are reasons that DEA would add no valuable insights.

Readers who sufficiently understand DEA or who prefer to return to this at a time when this more complex example can be more beneficial are encouraged to proceed to the subsequent section of this chapter describing how management can apply DEA.

The output and input data for the more complex two input-two output example are listed in Table 3-4.

Table 3-4. Two Input-Two output Example

Service Unit	Input 1	Input 2	Output 1	Output 2
S1	8.4	2.3	50	40
S2	9.0	3.2	40	50
S3	8.0	4.0	40	40
S4	9.0	2.0	30	50
S5	8.6	2.3	70	10
S6	4.8	2.2	20	30
S7	5.4	1.4	20	40
S8	9.8	4.8	40	80
S9	10.6	3.0	60	10
S10	10.2	3.2	60	20
S11	10.0	2.9	45	75
S12	4.5	2.5	10	30
S13	15.2	3.1	90	20
S14	11.8	4.3	90	30
S15	9.2	4.0	20	70
S16	9.2	2.5	70	20

These data were developed by arbitrarily assigning output levels to each service unit. An efficient amount of each input was then established based on an arbitrary formula and applied uniformly to each service unit. Hence, each service unit has a theoretically efficient input level to produce its outputs. These inputs were then increased arbitrarily to reflect inefficiencies in the operation. Several use more of one or both inputs to produce their outputs. While this example is not modeled on specific outputs or inputs, the outputs could be two types of purchase contracts and the inputs could be person years of buyers and clerical staff. Readers may wish to think in terms of other applications to more closely relate this analysis to their particular service activity.

The first question to consider is whether you can observe the inefficiencies in these service units. What approach would you use if you want to identify units that should be performing better; i.e., at lower cost or at greater output? Are there profit measures, historical standards, ratios, or regression analyses in use that direct management to problem areas and help develop programs to improve productivity and monitor that progress? Can we supply data in a way that makes the comparison more easily analyzed?

The reader is encouraged to consider the approach to this Table 3-4 and the complexity associated with just two inputs and two outputs. What if there were more units and more inputs and outputs?

Table 3-5. Two Input-Two Output Example: Sorted by Cost per Output

Service Unit	Output 1	Output 2	Total Output	% Excess Input 1	% Excess Input 2	Cost per Output	DEA rating
S11	45	75	120	0.00%	0.00%	2856	1
S7	20	40	60	0.00%	0.00%	3050	1
S8	40	80	120	13.79%	8.20%	3050	1
S1	50	40	90	0.00%	0.00%	3183	1
S16	70	20	90	0.00%	0.00%	3483	1
S14	90	30	120	6.21%	3.58%	3488	1
S2	40	50	90	16.30%	14.15%	3533	0.86
S6	20	30	50	14.29%	8.47%	3540	0.88
				MEDIAN			
S5	70	10	80	0.00%	0.00%	3656	1
S15	20	70	90	30.30%	26.79%	3733	0.95
S4	30	50	80	19.09%	20.50%	3750	0.95
S3	40	40	80	21.67%	16.00%	3750	0.88
S12	10	30	40	32.86%	26.09%	4313	0.82
S10	60	20	80	22.39%	21.19%	4425	0.78
S13	90	20	110	21.86%	23.88%	4568	1
S9	60	10	70	29.41%	28.93%	5186	0.70

*Actual cost and FTE's compared with efficient input level

Table 3-5 sorts these in order of size based on the cost per output. This was calculated by converting the input units into the total cost. These costs are incurred jointly to produce the outputs. The outputs were then simply added together to get the combined units and divided to get the cost per unit. This cost does not reflect the mix of outputs and the fact that one is more costly and resource intensive than the other. It is like combining appendectomy and cardiac bypass operations to measure the operating room output as number of operations. The result is the type of cost per unit that would be possible from an accounting system.

Two other columns that relate to the true inefficiency present in this hypothetical data set are provided. The "% Excess input 1" column reflects a comparison of the actual amount of excess input 1 compared with the underlying model of efficient performance. For example, S11, the first unit in Table 3-5 is using no excess input 1 while S9, the last unit is using over 29% excess input 1. Similarly the column titled "% Excess input 2" indicates the S11 is using the efficient amount of input 2 while S9 is using almost 29% excess of input 2 and S6 is using about 8.5% excess of input 2.

After reviewing table 3.5 and developing your own conclusions about how you might interpret these data, consider the following observations and comments about the relative strengths and limitations of DEA vis a vis other analytic methods.

1. Cost per unit ratio is a common type of financial measure and could generate the following insights:
   a. Accurately identifies many of the low cost units, such as S11, S7 and S1, as being the ones that are more efficient.
   b. Highest cost units, such as S9, S10 and S12, are among the least efficient.
   c. Suggests that unit S5, which is efficient, is in the higher cost half of the units suggesting it is inefficient. This misclassification could lead to requiring a completely efficient unit to reduce its resource level, which may be impossible without sacrificing other dimensions such as quality.
   d. What is the ideal cost per unit and how much cost savings and reduction in inputs is possible and reasonable to make the less efficient units as efficient as the best practice units? The best practice cost per unit is not provided by the ratio. Some average of the lower cost units might be selected or the lowest units might be selected as the benchmark. This would be somewhat arbitrary and, as in the above item c, it would pressure units such as S5, S16, and S1 to reduce their inputs even though they are not using excess resources and are relatively efficient.
2. One might create a cutoff point like the median to separate efficient and inefficient units.
   a. The cutoff of median or mean is arbitrary and has the same potential weaknesses as item 1d above. There is a mixture of efficient and inefficient units in the low cost group of units above the median line. How do you select the group to serve as the benchmark – median, mean, top quarter, top one? Any cutoff will be arbitrary and may create an artificial and unattainable goal for some of the other unit, such as S5.
3. DEA rating:
   a. Key attribute is that no units identified as inefficient are efficient. This means that DEA will not lead a manager to challenge an efficient unit to reduce resources when the unit is already using the minimum

needed to provide its volume and mix of services. This is the DEA foundation that allows mangers to use the analysis with confidence and without the risk of damaging their credibility from making this type of error.

b. Please recall two basic concepts described at the beginning of Chapter 2.
   i. If an efficient standard existed and was known, one would not need to calculate cost per unit or use DEA. We would know the exact amount of each resource needed and could directly determine whether a unit was inefficient and the amount of excess resources used.
   ii. Services rarely have efficient standards. Many organizations that have used DEA have something they would describe as a standard for some of their activities. DEA located substantial real inefficiencies beyond those identified by the standards because the standards were not really *efficient* standards. If the reader believes their service organization has efficient standards, common in services like fast foods, then DEA may provide less incremental benefits. Services where there is professional judgment such as lending, financial transaction processing, medical procedures, airplane maintenance, nursing care, tend to have no reliable efficient standard or incomplete standards. As will be illustrated in the application chapters, DEA offers one way to help develop these standards by using best practice units as the model for standard setting.

c. The efficiency rating indicates roughly the amount of resource saving that DEA has identified, but this amount may also understate the true inefficiency. This relates to the linear program structure that tries to make each unit as efficient as possible, giving each the benefit of the doubt. We also do not know where the true frontier is. As a result, the actual inefficiencies may be greater, and in practice, this has led to managers achieving greater savings that were suggested by DEA. For example, the efficient rating of

S9 is about 70%, suggesting use of about 30% excess resources. The excess of inputs 1 and 2 in Table 3-5 are also about 30%. In contrast, DEA underestimates the inefficiency in S15 and S4 where the actual inefficiency is 20% to 30% while the DEA rating is 95% suggesting only about 5% inefficiency.

d. DEA identifies specific amounts of excess inputs 1 and 2 and specific best practice model units, the ERS. This suggests specific paths to reduce resources by adopting ERS operating methods. Recall that ratios provide no details or suggestion as to what is best practice and do not specifically suggest the excess resources being used. Similarly, two of the units with inefficiencies identified by DEA, S2 and S6, might be considered among the better performing units based on cost ratios – they are in the upper half of the list suggest one of the lower cost half of units. Here and in business applications, DEA can and has located units that appear to be well managed and very profitable to indicate ways that their performance could be further improved to make them even more profitable.
e. The excess resources identified with DEA represent inefficient use of inputs which can be eliminated based on the ability of other service units to operate effectively with fewer inputs. Consequently, these resource reductions in inefficient units are likely to be sustainable and will not damage their operating effectiveness if they adopt the methods of the ERS best practice units related to the inefficient unit. This result is in contrast to cost cutting initiatives where across-the-board cuts are imposed resulting in already efficient units reducing resource levels that can damage operating effectiveness (quality, timeliness, etc.) and which will not be sustainable when the resulting operating weaknesses become visible.

Examples of the kinds of questions that may arise are as follows:

Can DEA help to locate real inefficiencies? Does it identify inefficiencies that are not real? To what extent does it fail to identify real inefficiencies present? Are other less complex approaches equally capable of defining ways to improve productivity? Are there approaches more powerful in their ability to locate inefficiencies and ways to improve productivity? Are there ways to coordinate these approaches to best achieve these results?

3.4. INTERPRETING DEA RESULTS

The DEA results for one of these two input-two output service units will be reviewed to relate the interpretation to the simpler examples already discussed. Note that as we approach more realistic service examples, we can no longer illustrate performance on a two-dimensional graph.

Service unit S12 is found to be inefficient by DEA with an efficiency rating of 0.817. S12 is also noted to be inefficient compared with S8, its efficiency reference set (See Figure 4-9 in chapter 4 for an example of the DEA results using Excel for this data set). While this may not have been apparent from initial observation, closer scrutiny allows one to observe this. S12 produces less than half of the outputs of S8 and uses about half of the inputs of S12. The coefficient assigned to S8 in comparing it to S12 is 0.375, which can be used to compare the adjusted S8 (previously referred to as the composite) with S12 based upon (4.2).

Figure 4-12 in Chapter 4 indicates that if S12 were as efficient as S8, it should be able to produce five added units of O2 while using 0.83 less units of I1, and 0.70 less units of I2. If we observe the u and v values assigned to S12, the greater weights are assigned to O2 and I1, which means that S12 would appear most efficient in its production of O2 and is using I1 more efficiently than I2. That is, a greater weight on the outputs increases the efficiency rating and a smaller weight on the inputs also would increase the efficiency rating. Recall that DEA seeks the weights that give the unit the highest efficiency rating possible subject to the constraint that all units cannot be more than 100% efficient when these same weights are applied.

Similar analyses can be conducted for each of the inefficient units. Note that 9 of these 16 service units are found to be inefficient. These are the units that management would focus on to improve productivity.

The efficiency rating does not really provide a basis for pure rank ordering of the most to least inefficient unit. Indeed, a small unit with high inefficiency may have potential less absolute productivity gains than a large unit with lower inefficiency. Technically, one can conclude that two units with the same efficiency reference set (ERS) can be ranked by the efficiency rating. For example, S2's rating of 0.86 means it is less efficient than S3 with an efficiency rating of 0.88 because they both have the same ERS . S12 cannot be ranked as more or less efficient than S2 or S3 because S12 has a different ERS.

Generally, management can assess the units that will be likely to yield important productivity gains by considering whether an increase in the efficiency rating to 1.0 would lead to substantial resource saving or output expansion. Here, an efficiency rating of 0.70 for S9 suggests 30% inefficiency is present, which would tend to be of more immediate concern than the inefficiency of 5% noted in S4. However, both units have potential productivity improvements and indeed, the actual improvements that will ultimately be achieved may be greater in unit S4 than S9.

Other models of DEA and further analysis of the results are useful in some applications. Some of these added capabilities are utilized in the application to the purchasing offices of the Government of Canada as is noted in Chapter 9 and in banking using a quality adjusted DEA model in chapter 7.

In addition to the static one-period analysis (which was used in the chapter illustrations), DEA can monitor and thereby help control and manage operating efficiency over time. DEA can be run with multiple-period information (e.g., quarters and years) for individual organization units or for each of a set of units to determine if the units are becoming more or less efficient compared to other units and to themselves over time. The use of DEA for successive periods can indicate whether previously inefficient units have become relatively efficient through remedial actions, and it can help locate other units that have become relatively inefficient.

3.5. REVIEW OF THE CAPABILITIES AND LIMITATIONS OF DEA

The conclusions to be derived regarding the types of insights provided by DEA are described below and are related back to the auto efficiency discussed at the beginning of Chapter 2.

1. We still do not know if the autos in our example (or the relatively efficient service units) could be made more efficient because we are comparing actual units against other actual units rather than comparing against an efficient standard.

2. We are able to identify inefficient service units and inefficient car usage by explicitly considering the mix and use of inputs to produce outputs.

3. If we use a dollar input measure, the inefficiency may be due to the price of supplies and tellers for the bank or the price of gasoline for the autos. If we use physical input measures, the inefficiency will not be due to price. However, we can run the DEA analysis two times, once with physical and once with dollar measures as one way of finding out whether price is a component of the identified inefficiency. For example, comparing the dollar and physical results, any units that are efficient with respect to physical units but inefficient with dollar measures are incurring higher costs per input but are not using excess inputs compared with the other units in the data set.

   Our preference is to analyze price efficiency more directly rather than rely on DEA for the analysis. This might be done by simply listing all inputs and input prices for each unit and analyzing these inputs one-by-one to determine if higher prices are due to controllable factors such as poor purchasing procedures or uncontrollable factors such as geographic location price differentials. Based on this analysis, actions to increase price efficiency can be developed and implemented.

4. After price is factored out of the analysis, the inefficiency may be due to scale or mix of inputs and outputs. While there are several ways to segregate these elements, we cannot totally do so with DEA alone. We can, however, determine when real inefficiencies are present, and investigation of the nature of these inefficiencies can lead to action to improve efficiency. For the auto example, one could require use of the more economical cars. For bank branches, one might decide to reduce staff, change transaction-processing procedures, and/or adjust the mix of inputs. We may also determine that scales economies are the source of the inefficiencies leading to a different set of appropriate remedies to improve productivity.

5. While DEA is unlikely to locate all inefficiencies, we know that the inefficiencies identified with DEA are sensitive to the mix of inputs and outputs and that they are real inefficiencies. Referring back to the auto example DEA eliminates the risk of erroneously concluding that Mary is inefficient in use of her autos and a tune up

is needed: rather we are more likely to identify Mary as price inefficient and John as inefficient with respect to mix of inputs used.

3.6. HOW MANAGEMENT CAN APPLY DEA

The process of applying DEA is summarized below. (This process is illustrated in coordination with use of other techniques in Chapter 6 for banking and Chapter 9 for government services.)

Step 1

Management should identify the units for which a DEA efficiency evaluation would be of interest. Generally this would be a set of units providing similar services for which management wants to evaluate performance and improve operating efficiency.

Step 2

Management should identify the relevant outputs and inputs of the units to be evaluated, and they should be measured for a representative period of time (e.g., a year, a quarter, a month). The relevant outputs are those services and other activities that the unit is responsible for in order to achieve its goals and objectives. The inputs are those resources that are required to produce the designated outputs.

Field applications of DEA have indicated that this process of output and input identification in itself is often useful to managers. The outputs and inputs are frequently not explicitly identified or understood. In addition, some of the relevant outputs and inputs may not have been measured or captured in the organization's management information system. The absence of data on relevant outputs and inputs has tended to raise questions about the adequacy of the information system, since these input and output data are needed to assess operating performance regardless of the technique that may be used. The outputs identified should generally be related to the inputs selected as follows: an efficient unit should respond to an increase or decrease in outputs with a corresponding increase or decrease in inputs.

If all the relevant outputs and inputs are not included in the DEA analysis, the DEA results will have to be reviewed for any bias that might result. For example, a DEA application to hospitals excluded a measure of the quality of services. Such use of DEA requires that the results be reconsidered to determine if the inefficient hospitals quality of care exceeds that of the efficient hospitals' by a large enough margin to compensate for the inefficiency identified with DEA. In this case it was found that quality of care was not a compensating factor. The results of other applications of DEA may, however, require some qualification if certain relevant input or output measures are excluded.

Specifying the input and output model is the most critical aspect of this process. Misspecification can destroy any value to this analysis. The examples in applications that follow in this book will, by example, suggest the way the model should be specified. In short, management needs to tenaciously ask are all the relevant inputs and outputs included and are they measured in a reasonably accurate manner. If similar services require very different amounts of effort and activity, they should be treated as two different outputs. Here, managements, judgment is critical to defining the outputs and identifying appropriate ways to measure the outputs.

As the number of outputs and inputs increase, the power of DEA decreases, to some degree. Similarly, as the number of service units decrease, the power of DEA to identify inefficiency also decreases. For some applications, this accurate characteristic of DEA has been used to suggest the minimum number of service units required and to justify using fewer inputs and outputs. Any manager planning to use DEA to make decisions that will change their business methods including changing the design of the unit and reducing the personnel in a unit needs to be sure that the analysis is as complete and reliable as possible.

While more service units make the results more distinctive in locating inefficiencies, DEA has been effective in service unit sets as low as seven and applications in government and financial services have proven valuable with fewer than 40 units. Consequently, our recommendation is to try the analysis with the number of service units and determine if the insights are valuable based on the analysis of the output. The effort to do this analysis is modest and the potential benefits are substantial enough to justify this even if the conclusion is that the insights available from DEA are not valuable. Some methods to enhance the power of DEA in these situations are discussed in the following chapter.

We also suggest that the number of inputs and outputs include all the ones the manager considers relevant and measurable. *Excluding relevant outputs to boost the power of DEA in cases where managers will act on this information can be dysfunctional.* Unrealistic models can lead to misleading information and decisions based on these incomplete models can generate costly errors. In some cases, when incomplete models adopted, scrutiny by more senior managers raised questions about whether personnel completing the analysis really understood their business.

Where relevant measures are excluded for whatever reason, the analysis naturally needs to be interpreted in light of any bias or erroneous conclusions that might result from the missing elements. Quality is one element that has been excluded in many basic DEA analyses, yet one that is critical for many service organizations. Methods of including quality in the analysis are described in the subsequent chapters.

Step 3

DEA is applied to the output and input data, and the results are analyzed to help management locate and remedy operating inefficiencies. Generally management will not have seen results similar to those derived from DEA, and these results will tend to provide insights not available from other widely used analytical techniques, such as ratio analysis. Management might begin by considering whether the location and magnitude of inefficiencies are consistent with their prior view of the operations of the service units being evaluated. This may raise questions about how complete and representative the output and input data are.

Step 4

The inefficient units should then be further studied and compared with their efficiency reference set units in order to ascertain the cause and controllability of the identified inefficiencies. In some cases, the inefficiencies may represent slack intentionally built into a unit or special circumstances that do not permit improvements in operating efficiency. In such circumstances, DEA is useful in promoting understanding the cost of this inefficiency and no further managerial action may be warranted. When the inefficiencies are found to be associated with the systems and managerial techniques used in these units, remedial action to improve efficiency should be implemented.

Insights from DEA will direct management's attention to aspects of the operations that are highly likely to benefit from remedial action. In contrast to other techniques, DEA evaluates units by explicitly and simultaneously considering the multiple inputs used to produce multiple outputs and without the need to know the efficient input/output relationships. Although DEA does not actually specify the remedial action needed, it focuses management's investigation on the inefficient units and their efficiency reference set. Through this process, DEA can be used to allocate management support to areas where weaknesses are known to exist and to help management identify ways in which management techniques can and should be improved. DEA provides management with a variety of paths to improve performance and management selects the paths that are deemed most feasible from a political, social and technical perspective. This decision process is described in applications covered in subsequent chapters.

3.7. OUTPUT-ORIENTED DEA MODEL

In general, the task of managing a service organization is so complex that refinements in the DEA model will only add marginally to the productivity improvements that may be achieved with the version described in this chapter. Some of these refinements may be justified solely because they incorporate more specific knowledge of the underlying service production relationships of outputs to inputs. In this case, the result may still be only marginally improved but the credibility and ultimate acceptance of the results may be enhanced and the DEA result may be more easily translated into operating changes to improve productivity. In some cases, the power of DEA to locate inefficient units will be improved by incorporating other data into a more elaborate model. This can increase the value of the information by locating more productivity improvement opportunities than are apparent from the original formulation. If DEA was applied and interpreted carefully in the basic form described above, the more elaborate version will not reverse any conclusions about where inefficiencies may be present but may increase the units identified as inefficient beyond the inefficiencies identified with the simple model pointing to greater potential resource savings that was previously identified with the basic DEA model. (One exception is the use of the variable returns to scale DEA model discussed in the following chapter. In anticipation of the discussion of returns to scale in DEA models, note that in Figure 3-4 there is a returns to scale column (RTS) which suggests for each service whether there is increasing, constant or decreasing returns to scale. In this artificial data set, the underlying data are constructed with constant returns to scale. Hence the meaning of these terms vis a vis the true behavior of the service organization needs to defined and the implications will be explained in chapter 5.) These other versions of DEA are described in the next two chapters. In this section, we only present an output-oriented version of models (2.1) and (2.2), as it is also a basic model that is a natural alternative to the input-oriented model that we have focused on in all the discussion up to this point.

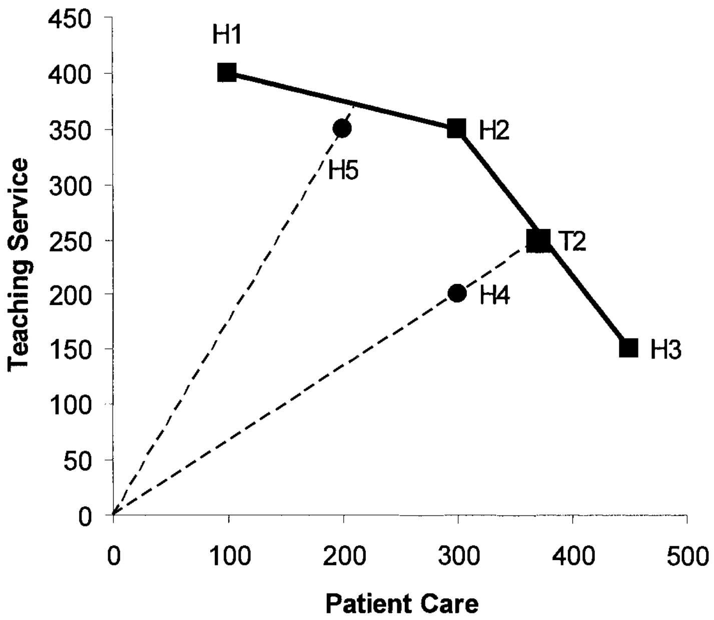


Figure 3-3. Five Hospitals

Models (2.1) and (2.2) in chapter 2 are called input-oriented DEA models, because they optimize the input usage while the outputs are fixed at their current levels. It is most appropriate in a situation where one can manage the resources used to provide service but where one cannot easily manage the service level or service demand. For example, a hospital can manage the resources it uses to care for those being treated, but it cannot influence how many emergency room visits, how many oncology cases it gets, etc. In cases where it can manage the service level, such as attracting elective surgery, the question may be how many more elective surgery cases can be handled with the current level of resources based on other more efficient hospitals' use of resources to provide elective surgery. Such changes would increase services and revenues in contrast to the previous examples focused on reducing resources and costs. This DEA analysis related to influencing the output volume and mix is referred to as an output-oriented model in contrast to the input-oriented models focusing reducing resources and costs.

Now, consider the five hospitals plotted in Figure 3-3. Each hospital provides a mix of patient care and teaching services as two outputs with the same amount of medical/surgical cost as the input. DEA identifies H1, H2 and H3 as the best practice units. H4 and H5 could increase their outputs with the current amount of input used compared to H1, H2 and H3. For example, the outputs produced by the linear combination of H2 and H3 (i.e., T2) can be larger than the current output levels of H4. The difference indicates the amount of inefficiency in H5 and implies the magnitude of added value (profits) when H5 achieves best practice level – T2.

To identify the DEA frontier portrayed in Figure 3-3, we use the following output-oriented (envelopment) model,

$$\max \phi + \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)$$

subject to

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} && i = 1,2,\ldots,m; && (3.3)\\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro} && r = 1,2,\ldots,s; &&\\
&\lambda_j \geq 0 && j = 1,2,\ldots,n. &&
\end{aligned}$$

where ε is a non-Archimedean which is defined to be less than any real positive number, ϕ represents the (output-oriented) efficiency score, and s_i^- and s_r^+ are input and output slacks, respectively.

The presence of the non-Archimedean ε in the objective function of (3.3) effectively allows the maximization over ϕ to preempt the optimization involving the slacks, s_i^- and s_r^+. As a result, Model (3.3) is also calculated in a two-stage process. First, we calculate ϕ^* by ignoring the slacks, and then we optimize the slacks by fixing the ϕ^* in the following linear programming problem.

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

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \qquad i = 1,2,\ldots,m; \tag{3.4}$$

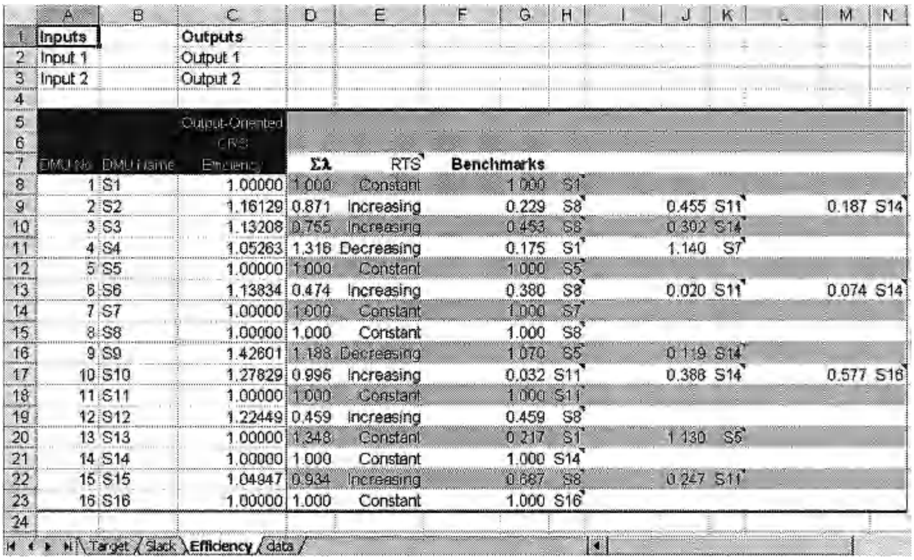
$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi^* y_{ro} \qquad r = 1,2,\ldots,s;$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n.$$

A service unit is efficient if and only if $\phi^* = 1$ and $s_i^{-*} = s_r^{+*} = 0$ for all i and r. The service unit is weakly efficient if $\phi^* = 1$ and $s_i^{-*} \neq 0$ and (or) $s_r^{+*} \neq 0$ for some i and r. If $\phi^* > 1$, then the service unit is inefficient.

Note that $\phi^* \geq 1$, and $\phi^* = 1$ if and only if $\theta^* = 1$. This indicates that models (1.5) and (1.6) identify the same frontier. Also, $\theta^* = 1/\phi^*$ (see, e.g., Zhu (2003)).

In short, the output model will identify the exact same units as inefficient as the input model. However, the focus is on increasing the output and this generates a somewhat different set of λ values.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Inputs		Outputs											
2	Input 1		Output 1											
3	Input 2		Output 2											
4														
5			Output-Oriented											
6			CRS											
7	DMU No.	DMU Name	Efficiency	Σλ	RTS	Benchmarks								
8	1	S1	1.00000	1.000	Constant		1.000	S1						
9	2	S2	1.16129	0.871	Increasing		0.229	S8		0.455	S11		0.187	S14
10	3	S3	1.13208	0.755	Increasing		0.453	S8		0.302	S14			
11	4	S4	1.05263	1.316	Decreasing		0.175	S1		1.140	S7			
12	5	S5	1.00000	1.000	Constant		1.000	S5						
13	6	S6	1.13834	0.474	Increasing		0.380	S8		0.020	S11		0.074	S14
14	7	S7	1.00000	1.000	Constant		1.000	S7						
15	8	S8	1.00000	1.000	Constant		1.000	S8						
16	9	S9	1.42601	1.188	Decreasing		1.070	S5		0.119	S14			
17	10	S10	1.27829	0.996	Increasing		0.032	S11		0.388	S14		0.577	S16
18	11	S11	1.00000	1.000	Constant		1.000	S11						
19	12	S12	1.22449	0.459	Increasing		0.459	S8						
20	13	S13	1.00000	1.348	Constant		0.217	S1		1.130	S5			
21	14	S14	1.00000	1.000	Constant		1.000	S14						
22	15	S15	1.04947	0.934	Increasing		0.687	S8		0.247	S11			
23	16	S16	1.00000	1.000	Constant		1.000	S16						
24														

Figure 3-4. Output-oriented DEA Efficiency

For example, if we apply model (3.4) to the data in Table 3-4, the output-oriented DEA efficiency scores are reported in column C of Figure 3-4[1]. (See file "Table 3-4 output oriented.xls" for the complete DEA results.) One can easily verify that $\theta^* = 1/\phi^*$.

The dual program to (3.3) is the multiplier version of the output-oriented DEA model,

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

subject to

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \quad j = 1,\dots,n \tag{3.5}$$

$$\sum_{r=1}^{s} u_r y_{ro} = 1$$

$$u_r, v_i \geq \varepsilon > 0$$

	A	B	C	D	E	F	G
1	Inputs		Outputs				
2	Input 1		Output 1				
3	Input 2		Output 2				
4							
5			Output-Oriented				
6			CRS	Optimal Multipliers			
7	DMU No.	DMU Name	Efficiency	Input 1	Input 2	Output 1	Output 2
8	1	S1	1.00000	0.00000	0.43478	0.01304	0.00870
9	2	S2	1.16129	0.12903	0.00000	0.01398	0.00882
10	3	S3	1.13208	0.14151	0.00000	0.01533	0.00967
11	4	S4	1.05263	0.00000	0.52632	0.01579	0.01053
12	5	S5	1.00000	0.10802	0.03086	0.01327	0.00710
13	6	S6	1.13834	0.23715	0.00000	0.02569	0.01621
14	7	S7	1.00000	0.16187	0.08993	0.00000	0.02500
15	8	S8	1.00000	0.10204	0.00000	0.01105	0.00697
16	9	S9	1.42601	0.13453	0.00000	0.01570	0.00583
17	10	S10	1.27829	0.11420	0.03544	0.01393	0.00822
18	11	S11	1.00000	0.10000	0.00000	0.01083	0.00683
19	12	S12	1.22449	0.27211	0.00000	0.00000	0.03333
20	13	S13	1.00000	0.00000	0.32258	0.00968	0.00645
21	14	S14	1.00000	0.08475	0.00000	0.00918	0.00579
22	15	S15	1.04947	0.09396	0.04027	0.00000	0.01429
23	16	S16	1.00000	0.10024	0.03111	0.01222	0.00721
24							

Target / Slack / Efficiency / **Efficiency Report** / data

Figure 3-5. Output-oriented DEA Efficiency and Optimal Multipliers

[1] These results are calculated using the *DEAFrontier* software included in the CD (see chapter 4 on how to use *DEAFrontier*.

If we apply the multiplier model (3.5) to the data in Table 3-4, the output-oriented DEA efficiency scores and a set of optimal multipliers are reported in the Efficiency Report sheet of the Excel solution worksheet, as shown in Figure 3-5. (See file "Table 3-4 output oriented.xls" for the complete DEA results.)

To illustrate the contrasting message of the input-oriented versus the output oriented DEA model, consider the unit S9, which had the lowest efficiency rating with the input-oriented model of 0.70 or 70% (see Table 3-5). The rating with the output-oriented model (Figure 3-4) is 1.426, which is more than 1.0 signifying that is an inefficient unit. The input-oriented rating of .70 is equal the one divided by the output oriented rating or 1.426. The multipliers for the input and output variables are the same. The efficiency reference group for S9 is also identical.

Table 3-6. Input- versus Output-oriented DEA models

	Input 1	Input 2	Output 1	Output 2
S9 – Actual inputs and outputs	10.6	3.0	60	10
Input oriented target	7.4	2.1	60	10
Actual – input-oriented target – resource reductions	3.2	0.9	0	0
Output-oriented target	10.6	3.0	85.6	14.2
Actual – output-oriented target – output increases	0	0	- 25.6	- 4.2

The key difference from a managerial perspective is the slack values and the excess resources or additional services that the models suggest would make S9 as efficient as its ERS. These are most directly reported in the table that describes the Target[2] inputs and outputs for each unit, and suggest what the input and output levels could be if the unit was performing as well as its ERS. The differences between these models, summarized by Table 3-6 are as follows: The input-oriented model suggests that S9 has the potential to reduce Input 1 by 3.2 units and Input 2 by 2.1 units to become as efficient as the best practice ERS units. The suggestion is that no change to the outputs would be possible. (The output change will not always be zero, and will in fact often suggest that a unit like S9 can achieve the above savings and also increase outputs a modest amount.) In contrast, the output-oriented model suggests that S9 has the potential to increase output 1 by 25.6 units and output 2 by 4.2 units with *no* reduction in inputs. A manager able to control their service level

[2] These Target values are reported by the *DEAFrontier* software (see chapter 4).

more they can control their resource level might find the output-model more valuable. Similarly, if one wanted to know how much S9 could grow before requiring added resources, this output-oriented model would be preferred. If the organization were seeking increased profits from cost reductions, the input-oriented model would be more valuable. As always, the actual change in inputs and outputs will be determined by management's assessment of the practicality and feasibility of the alternate paths to improve productivity.

3.8. CONCLUSION

Circumventing the DEA black box: The above description of DEA is intended to make this powerful technique more accessible to managers. Working with data sets that the reader is familiar with the Excel DEA program provided can provide familiarity and confidence in understanding and interpreting the results. There is one final view that has proven effective in relaying the value of the DEA analysis. Return to the 2-input 2-output data set in Table 3-4 and consider whether you could have identified the inefficient units and the amount of resource savings that were possible. DEA indicates that S9 has and ERS of S14 and S5. In many cases, at this point, one needs only put these units on a chart to see that S9 should be able to improve its performance.

Table 3-7. ERS for S9

	Service Unit	Input 1	Input 2	Output 1	Output 2
ERS for S9	S5	8.6	2.3	70	10
ERS for S9	S14	11.8	4.3	90	30
Inefficient unit	S9	10.6	3.0	60	10

Simply looking at the inputs of S9 compared with S5 (see Table 3-7); one can see that S5 produces the same or more of each output and uses fewer of each input than S9. Similarly, S14 produces 50% more output 1 and three times the amount of output 2 than S9 and uses about 11% more of input 1 and about 43% more of input 2. One could easily conclude that S9 should be able to use less resources compared with S5. DEA has enabled us to locate this comparison. When the number of units increases, this ability to find these otherwise invisible but convincing examples of best practice units providing services with fewer resources becomes much more remarkable. At this point, the existence of S5 can convince a manager of the validity of the inefficiencies located in S14 without the need to understand anything about the way DEA determines this and anything about the underlying linear programming[3].

DEA has proved useful in managing service productivity when applied as described in this chapter. It complements other common methods of measuring performance such as ratio analysis and has advantages over other techniques in its sensitivity to multiple outputs and inputs, and its objective location of inefficient units without the need for relative weights or value.

[3] The question of whether S 14 is as convincing would depend on managements' interpretation of the results with their perspective on the business and the tradeoffs between various types of inputs and service outputs.

Chapter 4

SOLVING DEA USING *DEAFRONTIER* SOFTWARE

4.1. INTRODUCTION

One can solve DEA models (e.g., (2.1) and (2.2)) using the spreadsheets and Excel Solver as described in chapter 2 or Zhu (2003). In this chapter, we demonstrate how to solve models (2.1) and (2.2) using the *DEAFrontier* software supplied with the book. The next chapter demonstrates the use of *DEAFrontier* to solve for other DEA models[1].

In order to use the DEAFrontier, please set the Macro Security to Medium Level or lower in the Excel. This can be done by selecting the Tools/Options menu item. In the Option menu item, click the Macro Security button and then select the "Medium" option as shown in Figures 4-1 and 4-2.

[1] See Zhu (2003) or www.deafrontier.com for other versions of the *DEAFrontier* software.

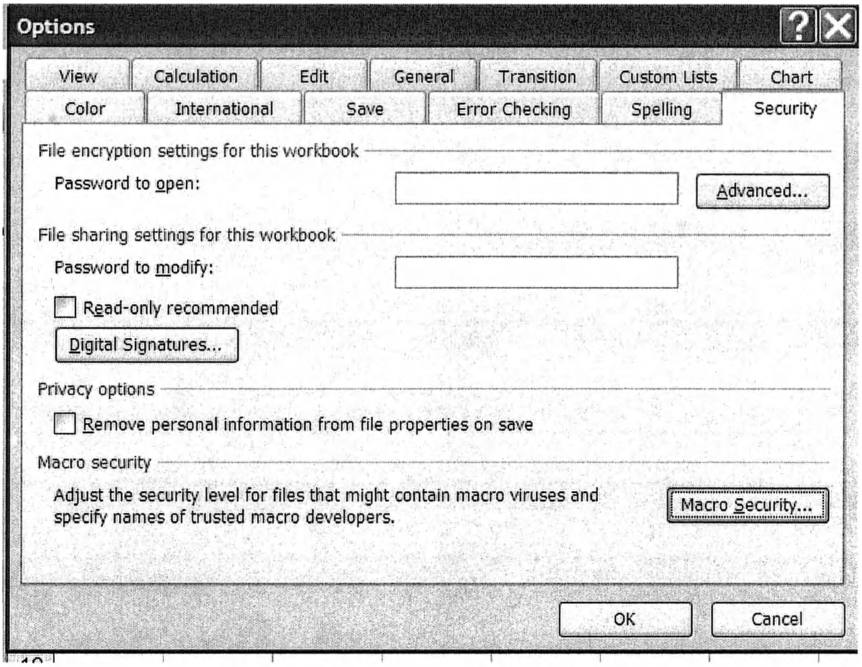


Figure 4-1. Macro Security

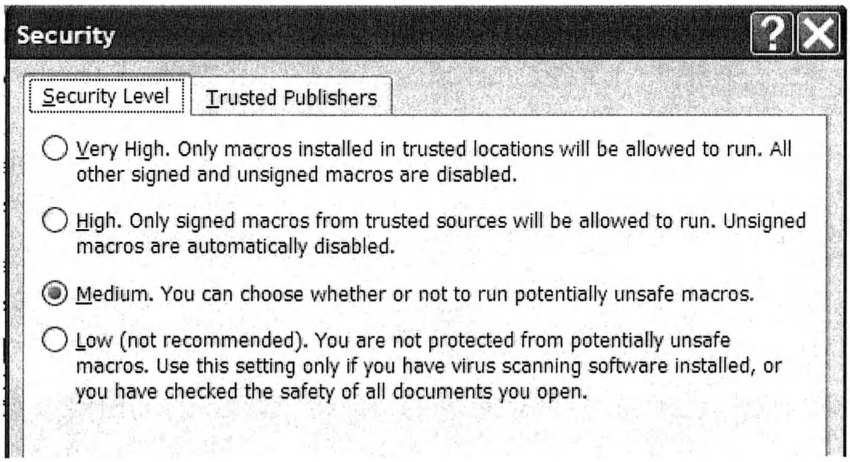


Figure 4-2. Medium Security Level

4.2. DEAFRONTIER

DEAFrontier is an Add-In for Microsoft® Excel and uses the Excel Solver. This software requires Excel 97 or later versions. The formulation is similar to the program defined in Chapter 2 that guided the reader in using their Excel program with the solver to complete a DEA analysis. While users will ultimately use the software provided or another DEA code to complete these analyses, a perspective on what is being done provides added confidence and the ability to interpret the results. At the same time, one need not be aware of the underlying model as long as the application uses an appropriate data set and the interpretation is consistent with the data and the power of the DEA methodology. Few who drive cars or use personal computers understand how they work, but can use them effectively as long as they know how to use these devices and appreciate their limitations.

To install the software the CD-ROM using Windows, you may follow these steps:

Step 1. Insert the CD-ROM into your computer's CD-ROM drive.
Step 2. Launch Windows Explore.
Step 3. Click Browse to browse the CD and find the file "*DEAFrontier.xla*". Copy this file to your hard drive. (see footnote 2.)

DEAFrontier does not set any limit on the number of units, inputs or outputs. However, please check www.solver.com for problem sizes that various versions of Excel Solver can handle (see Table 4-1).

Table 4-1. Microsoft® Excel Solver Problem Size

Problem Size:	Standard Excel Solver	Premium Solver	Premium Solver Platform
Variables x Constraints	200 x 200	1000 x 8000	2000 x 8000

Source: www.solver.com

To run *DEAFrontier*, the Excel Solver must first be installed, and the Solver parameter dialog box must be displayed at least once in the Excel session[2]. Otherwise, an error may occur when you

[2] *To avoid this step, one can copy the file "DEAFrontier.xla" to the subdirectory where the Excel Solver is installed. The Excel Solver is usually installed at 1)*

run the software, as shown in Figure 4-3. (*Please also make sure that the Excel Solver works properly. One can use the file "solvertest.xls" to test whether the Excel Solver works. This test file is also available at www.deafrontier.com/solvertest.xls.*)

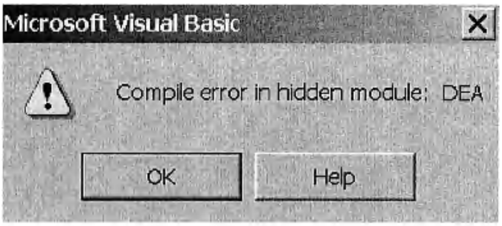


Figure 4-3. Error Message

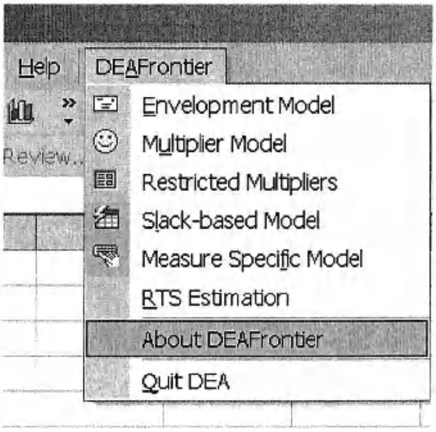


Figure 4-4. DEAFrontier Menu

You may follow the following steps.

First, in Excel, invoke the Solver by using the Tools/Solver menu item as shown in Figure 2-3 in chapter 2. This will load the Solver parameter dialog box as shown in Figure 2-5 in chapter 2. Then close the Solver parameter dialog box by clicking the Close button. Now, you have successfully loaded the Excel Solver.

Next, open the file DEAFrontier.xla, and a "DEAFrontier" menu is added at the end of the Excel menu. (see Figure 4-4). Now, the

---

Office XP: C:\Program Files\Microsoft Office\Office10\ Library\ Solver or 2) Office 2003: C:\Program Files\Microsoft Office\ Office11\Library\Solver)

DEAFrontier software is ready to run. *NOTE – Each time you reopen the Excel program, you will need reopen the DEAFronties.xla to make* the file ready for DEA computations.

4.3. ORGANIZE THE DATA

The data sheet containing the data for service units (or DMUs – decision making units) under evaluations must be named as "Data". The data sheet should have the format as shown in Figure 4-5. These instructions for organizing the data in the required cells of a spreadsheet labeled "Data" need to be followed very carefully to enable the DEASolver Excel program to locate the data and complete the DEA calculation.

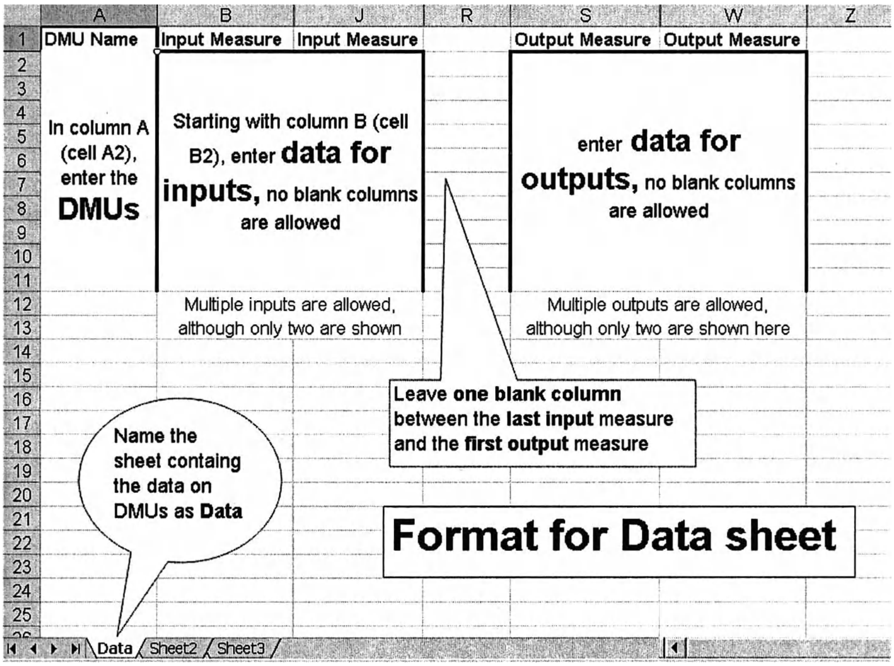


Figure 4-5. Data Sheet Format

Decision making units or DMUs are the most common term used in the DEA literature to refer to the units being evaluated. While the service units and providers that are the focus of this book are often decision making units, others are the subject of decisions that are made by external managers about their operations. While we could use the DMU designation, and the terms would be fully interchangeable, we have referred to the units being evaluated as service units or service providers and no technical distinction is intended.

Leave one blank column between the input and output data. No blank columns and rows are allowed within the input and output data. See Figure 4-6 for an example.

	A	B	C	D	E	F
1	Service Unit	Input 1	Input 2		Output 1	Output 2
2	S1	8.4	2.3		50	40
3	S2	9.0	3.2		40	50
4	S3	8.0	4.0		40	40
5	S4	9.0	2.0		30	50
6	S5	8.6	2.3		70	10
7	S6	4.8	2.2		20	30
8	S7	5.4	1.4		20	40
9	S8	9.8	4.8		40	80
10	S9	10.6	3.0		60	10
11	S10	10.2	3.2		60	20
12	S11	10.0	2.9		45	75
13	S12	4.5	2.5		10	30
14	S13	15.2	3.1		90	20
15	S14	11.8	4.3		90	30
16	S15	9.2	4.0		20	70
17	S16	9.2	2.5		70	20

data / Sheet2 / Sheet3

Figure 4-6. Sample Data Sheet

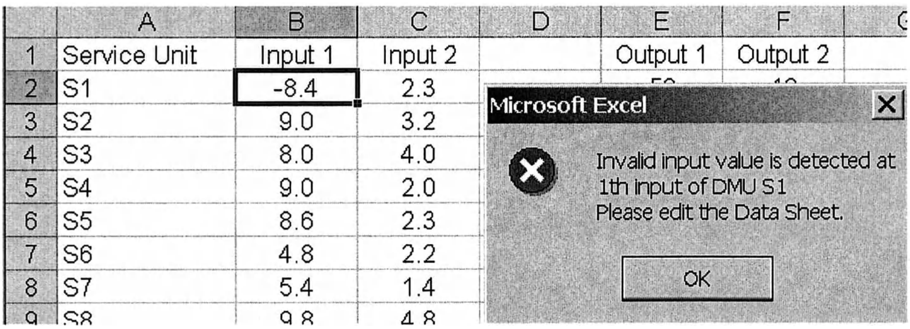


Figure 4-7. Invalid Data

Negative or non-numerical data are deemed as invalid data. The software checks if the data are in valid form before the calculation. If the data sheet contains negative or non-numerical data, the software will quit and locate the invalid data cell(s) (see Figure 4-7).

4.4. RUN THE *DEAFRONTIER* SOFTWARE

To run the envelopment model (2.2), select the "Envelopment Model" menu item. You will be prompted with a form for selecting the models, as shown in Figure 4-8.

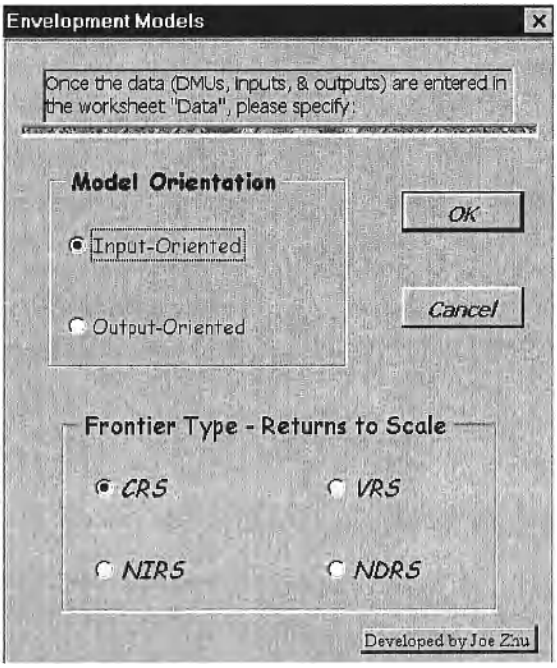


Figure 4-8. Envelopment Model

Model Orientation refers to whether a DEA model is input-oriented or output-oriented and Frontier Type refers to the returns to scale type of the DEA efficient frontier. We will discuss these concepts in chapter 5. Model (2.2) is an input-oriented DEA model whose frontier exhibits CRS (constant returns to scale). Therefore, we use the default selection specified by the *DEAFrontier*. (If we select the "Output-Oriented", then we will run the output-oriented CRS envelopment model (3.3) in chapter 3.)

The software performs a two-stage DEA calculation. First, the efficiency scores are calculated by using model (2.2), and the efficiency scores and benchmarks (Efficiency Reference Set) (λ_j^*) are reported in the "Efficiency" sheet, as shown in Figure 4-9.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Inputs		Outputs											
2	Input 1		Output 1											
3	Input 2		Output 2											
4														
5			Input-Oriented											
6			CRS											
7	DMU No.	DMU Name	Efficiency	$\Sigma\lambda$	RTS	Benchmarks								
8	1	S1	1.00000	1.000	Constant		1.000	S1						
9	2	S2	0.86111	0.750	Increasing		0.197	S8		0.392	S11		0.161	S14
10	3	S3	0.88333	0.667	Increasing		0.400	S8		0.267	S14			
11	4	S4	0.95000	1.250	Decreasing		0.167	S1		1.083	S7			
12	5	S5	1.00000	1.000	Constant		1.000	S5						
13	6	S6	0.87847	0.417	Increasing		0.334	S8		0.018	S11		0.065	S14
14	7	S7	1.00000	1.000	Constant		1.000	S7						
15	8	S8	1.00000	1.000	Constant		1.000	S8						
16	9	S9	0.70126	0.833	Increasing		0.750	S5		0.083	S14			
17	10	S10	0.78230	0.779	Increasing		0.025	S11		0.303	S14		0.451	S16
18	11	S11	1.00000	1.000	Constant		1.000	S11						
19	12	S12	0.81667	0.375	Increasing		0.375	S8						
20	13	S13	1.00000	1.348	Constant		0.217	S1		1.130	S5			
21	14	S14	1.00000	1.000	Constant		1.000	S14						
22	15	S15	0.95286	0.890	Increasing		0.654	S8		0.235	S11			
23	16	S16	1.00000	1.000	Constant		1.000	S16						
24														

Target / Slack / Efficiency / data

Figure 4-9. Efficiency Sheet (Envelopment Model)

The "Efficiency" sheet reports the input and output names. Column A reports the DMU No., where DMU stands for "decision making units". Column B reports the service units. Column C reports the efficiency scores (it also indicates the type of DEA models used). Column D reports the optimal $\sum\lambda_j^*$ which is used to identify the returns to scale classifications reported in column E (see next chapter for discussions on returns to scale). The Efficiency Reference Set is reported under the "Benchmarks".

At the same time, a "Slack" sheet is generated based upon the efficiency scores and the λ_j^* using the following formula

$$\begin{cases} s_i^- = \theta^* x_{io} - \sum_{j=1}^{n} \lambda_j x_{ij} & i = 1,2,\ldots,m \\ s_r^+ = \sum_{j=1}^{n} \lambda_j y_{rj} - y_{ro} & r = 1,2,\ldots,s \end{cases} \tag{4.1}$$

Then a "Target" sheet is generated based upon (3.2) in chapter 3, but the slacks are calculated from (4.1).

Note that the slacks calculated from (4.1) are not optimized and do not necessarily reflect the DEA slack. For example, consider B8 in Figure 3-1. $\theta^* = 1$, $\lambda^*_{B8} = 1$ and all other $\lambda^*_j = 0$ is a set of optimal solutions from model (2.2), indicating that all the slacks are equal to zero based upon (4.1). However, as demonstrated previously, model (3.1) in chapter 3 indicates that $s_H^{-*} = 10$ and $s_S^{-*} = s_T^{+*} = 0$. Because of possible multiple optimal solutions, (4.1) may not yield all the non-zero slacks.

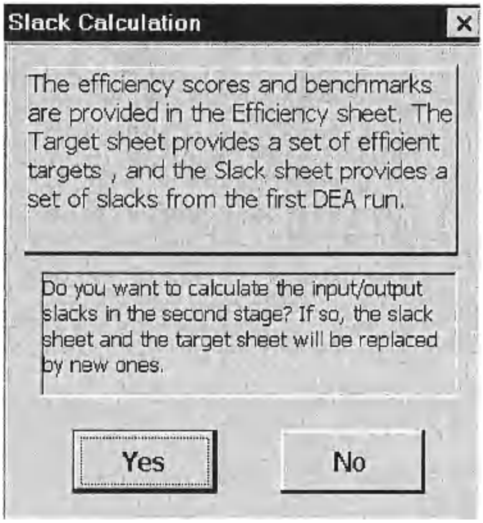


Figure 4-10. Slack Calculation

You will be asked whether you want to perform the second-stage calculation, i.e., fixing the efficiency scores and calculating the DEA slacks by using model (3.1) (see Figure 4-10). If Yes, then the Slack and Target sheets will be replaced by new ones based upon (3.1) and (3.2) (see Figures 4-11 and 4-12).

	A	B	C	D	E	F
1	**Inputs**		**Outputs**		Second Stage	
2	Input 1		Output 1			
3	Input 2		Output 2			
4						
5	Input-Oriented					
6	CRS Model Slacks					
7			*Input Slacks*		**Output Slacks**	
8	DMU No.	DMU Name	*Input 1*	*Input 2*	**Output 1**	**Output 2**
9	1	S1	0.13793	0.00000	0.00000	0.00000
10	2	S2	0.00000	0.41667	0.00000	0.00000
11	3	S3	0.00000	1.31111	0.00000	0.00000
12	4	S4	1.88333	0.00000	0.00000	0.00000
13	5	S5	0.00000	0.00000	0.00000	0.00000
14	6	S6	0.00000	0.70486	0.00000	0.00000
15	7	S7	0.00000	0.00000	0.00000	0.00000
16	8	S8	0.00000	0.00000	0.00000	0.00000
17	9	S9	0.00000	0.02044	0.00000	0.00000
18	10	S10	0.00000	0.00000	0.00000	0.00000
19	11	S11	0.00000	0.00000	0.00000	0.00000
20	12	S12	0.00000	0.24167	5.00000	0.00000
21	13	S13	3.71429	0.00000	0.00000	0.00000
22	14	S14	0.00000	0.00000	0.00000	0.00000
23	15	S15	0.00000	0.00000	16.76614	0.00000
24	16	S16	0.00000	0.00000	0.00000	0.00000

Target | **Slack** | Efficiency | Efficiency Report | data

Figure 4-11. Slack Sheet

In these two sheets, Cell E1 indicates whether the slacks and targets are obtained in the first or second stage DEA calculation. See file "ServiceChapter4.xls" for the DEA results.

The slacks in Figure 4-11 are calculated by DEA and are used with the efficiency rating to obtain the target values presented in Figure 4-12. For example, for S9, the target value = (actual input) X (efficiency rating)-slack. For I1, which has zero slack, the calculation is (10.6)(.701) – 0 = 7.43 the target for I1. For I2, which has slack of 0.02044, the calculation is (3.0)(.701)-0.24167 = 2.083, the target for I2. The difference between the target and the actual is the potential resource savings if the service unit operates as efficiently as the best practice ERS units.

	A	B	C	D	E	F	G
1	**Inputs**		**Outputs**		Second Stage		
2	Input 1		Output 1				
3	Input 2		Output 2				
4							
5	Input-Oriented						
6	CRS Model Target						
7			*Efficient Input Target*			**Efficient Output Target**	
8	***DMU No.***	***DMU Name***	*Input 1*	*Input 2*		**Output 1**	**Output 2**
9	1	S1	8.26207	2.30000		50.00000	40.00000
10	2	S2	7.75000	2.33889		40.00000	50.00000
11	3	S3	7.06667	2.22222		40.00000	40.00000
12	4	S4	6.66667	1.90000		30.00000	50.00000
13	5	S5	8.60000	2.30000		70.00000	10.00000
14	6	S6	4.21667	1.22778		20.00000	30.00000
15	7	S7	5.40000	1.40000		20.00000	40.00000
16	8	S8	9.80000	4.80000		40.00000	80.00000
17	9	S9	7.43333	2.08333		60.00000	10.00000
18	10	S10	7.97942	2.50335		60.00000	20.00000
19	11	S11	10.00000	2.85000		45.00000	75.00000
20	12	S12	3.67500	1.80000		15.00000	30.00000
21	13	S13	11.48571	3.10000		90.00000	20.00000
22	14	S14	11.80000	4.30000		90.00000	30.00000
23	15	S15	8.76633	3.81145		36.76614	70.00000
24	16	S16	9.20000	2.50000		70.00000	20.00000

Target / Slack / Efficiency / Efficiency Report / data

Figure 4-12. Target Sheet

Managers may find the most important single insight from DEA to be the content of Figure 4-12. The difference between the target and the actual input levels indicates the potential resource reductions (and cost savings) for each input based on the actual performance of other best practice units. All of the input reductions together would increase that unit's productivity to the best practice level. This information and the efficiency rating provide the unique insights that make DEA so valuable for service performance management.

To run the multiplier model (2.1), select the "Multiplier Model" menu item. You will be prompted with a form for selecting the models. The form is similar to the one shown in Figure 4-8. The results are reported in a sheet named "Efficiency Report", as shown in Figure 4-13 where the DEA efficiency and optimal multipliers are reported. The interpretation (and possible misinterpretation) of these multipliers is covered in the previous chapter in section 3.2.

	A	B	C	D	E	F	G
1	**Inputs**		**Outputs**				
2	Input 1		Output 1				
3	Input 2		Output 2				
4							
5			Input-Oriented				
6			CRS	Optimal Multipliers			
7	DMU No.	DMU Name	Efficiency	Input 1	Input 2	Output 1	Output 2
8	1	S1	1.00000	0.00000	0.43478	0.01304	0.00870
9	2	S2	0.86111	0.11111	0.00000	0.01204	0.00759
10	3	S3	0.88333	0.12500	0.00000	0.01354	0.00854
11	4	S4	0.95000	0.00000	0.50000	0.01500	0.01000
12	5	S5	1.00000	0.10802	0.03086	0.01327	0.00710
13	6	S6	0.87847	0.20833	0.00000	0.02257	0.01424
14	7	S7	1.00000	0.16187	0.08993	0.00000	0.02500
15	8	S8	1.00000	0.10204	0.00000	0.01105	0.00697
16	9	S9	0.70126	0.09434	0.00000	0.01101	0.00409
17	10	S10	0.78230	0.08934	0.02773	0.01090	0.00643
18	11	S11	1.00000	0.10000	0.00000	0.01083	0.00683
19	12	S12	0.81667	0.22222	0.00000	0.00000	0.02722
20	13	S13	1.00000	0.00000	0.32258	0.00968	0.00645
21	14	S14	1.00000	0.08475	0.00000	0.00918	0.00579
22	15	S15	0.95286	0.08953	0.04409	0.00000	0.01361
23	16	S16	1.00000	0.10024	0.03111	0.01222	0.00721
24							

Target / Slack / Efficiency / **Efficiency Report** / data

Figure 4-13. Efficiency Report Sheet (Multiplier Model)

Chapter 5

DEA MODEL - EXTENSIONS

Alternate models for specific service characteristics, to increase insights generated with DEA

5.1. INTRODUCTION

This chapter reviews a group of DEA models that were adapted to specific needs of service and other organizations. We have selected the models that are most likely to be needed and have noted the existence of other models that are primarily required for research with appropriate references. DEA continues to be adapted for new research purposes and the list of other models and substitutes for DEA noted in this chapter is not meant to be exhaustive. In addition, various groups have attributed different names and use somewhat different terms to refer to elements of the DEA models, which result in terminology differences that can be confusing. We do not presume to cover all the terminology differences and apologize in advance for any confusion caused by differing terminology. While we use a set of terms that are familiar to a large group of DEA researchers, we do not suggest that this is necessarily superior to other sets of labels and model names.

5.2. RETURNS TO SCALE FRONTIERS

The frontier determined by the DEA models discussed in chapters 3 and 4 exhibits constant returns to scale (CRS). This is sometimes rephrased to describe the basic model as reflecting piece-wise constant returns to scale. That is the rate of substitution between inputs and between outputs is constant within each segment of the frontier. This means each different segment will have somewhat different returns to scale characteristics. Some applications required explicit recognition that variable returns to scale might exist. Scale diseconomies can appear as an inefficiency using the basic DEA 1 but the inefficiencies may be caused by the size or the unit or volume of activity rather than use of excess resources due to poor management. For example, a small unit compared with a large unit might be classified as inefficient because it produces half the services of a large unit but uses 60% of the resources. The basic DEA model will accurately indicate that the small unit is inefficient compared with the large unit. However, if there are increasing returns to scale in this set of services, the inefficiency may be solely due to size so this inefficiency may not be one that can be eliminated without making the small unit a large one that can benefit from the scale economies in this business. The non-constant or variable returns to scale DEA model filters out the scale effects and would rate the small unit as efficient, avoiding the impression that is it inefficient. There are several important managerial issues that need to be considered in using this model versus the basic model that are discussed at the end of this section.

Returns to scale (RTS) have typically been defined only for single output situations. RTS are considered to be increasing if a proportional increase in all the inputs results in a more than proportional increase in the single output. Let α represent the proportional input increase and β represent the resulted proportional increase on the single output. Increasing returns to scale prevail if $\beta > \alpha$ and decreasing returns to scale prevail if $\beta < \alpha$.

Figure 5-1 shows a CRS frontier – ray OB. Based upon this CRS frontier, only B is efficient. In fact, the models discussed in the chapters 2 and 3 yield the efficient frontier that only exhibits CRS. Consequently, these models are called CRS (DEA) models.

In Figure 5-1, line segment AB exhibits increasing returns to scale (IRS) compared to OB, and BC and CD exhibit decreasing returns to scale (DRS) compared to OB.

Based upon the different types of RTS frontiers, we can establish different DEA envelopment and multiplier models.

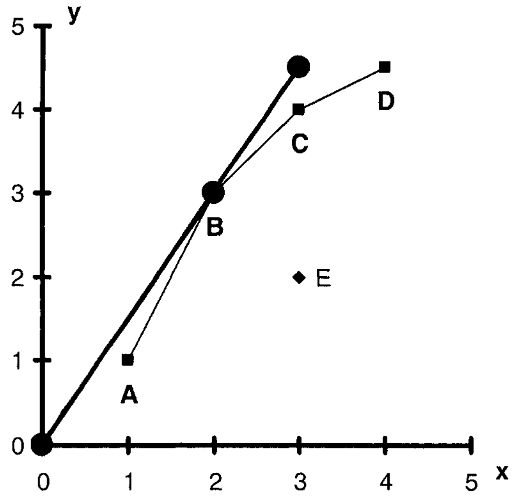


Figure 5-1. Returns to Scale Frontiers

5.3. NON-CONSTANT RETURNS TO SCALE DEA MODELS

The constraint on $\sum_{j=1}^{n} \lambda_j$ in the envelopment models actually determines the returns to scale (RTS) type of an efficient frontier. If we add $\sum_{j=1}^{n} \lambda_j = 1$, we obtain VRS (variable RTS) models. The frontier is ABCD as shown in Figure 5-1.

If we replace $\sum_{j=1}^{n} \lambda_j = 1$ with $\sum_{j=1}^{n} \lambda_j \leq 1$, then we obtain non-increasing RTS (NIRS) envelopment models. In Figure 5-2, the NIRS frontier consists of DMUs B, C, D and the origin.

If we replace $\sum_{j=1}^{n} \lambda_j = 1$ with $\sum_{j=1}^{n} \lambda_j \geq 1$, then we obtain non-decreasing RTS (NDRS) envelopment models. In Figure 5-3, the NDRS frontier consists of DMUs, A, B, and the section starting with B on ray OB.

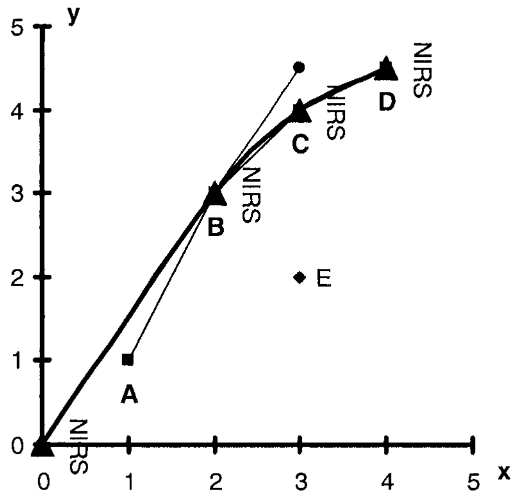


Figure 5-2. Non Increasing Returns to Scale (NIRS) Frontier

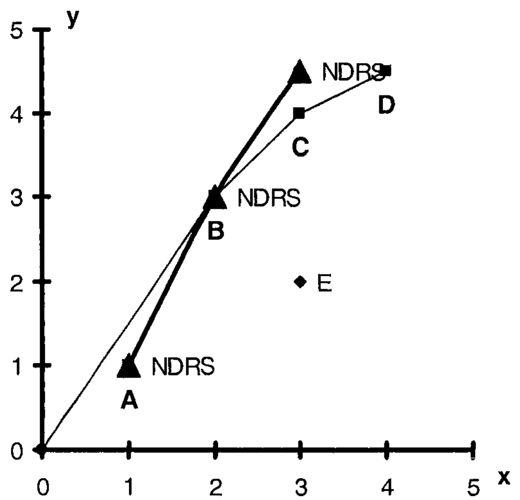


Figure 5-3. Non Decreasing Returns to Scale (NDRS) Frontier

Table 5-1 summarizes the envelopment models with respect to the orientations and frontier types. The last row labeled efficient tarhets presents the efficient target (DEA projection) of a specific service unit under evaluation.

Table 5-1. Envelopment Models

Frontier Type	Input-Oriented	Output-Oriented
CRS	$\min \theta - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io}$ $i = 1,2,...,m;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}$ $r = 1,2,...,s;$ $\lambda_j \geq 0$ $j = 1,2,...,n.$	$\max \phi + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}$ $i = 1,2,...,m;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro}$ $r = 1,2,...,s;$ $\lambda_j \geq 0$ $j = 1,2,...,n.$
VRS	Add $\sum_{j=1}^{n} \lambda_j = 1$	
NIRS	Add $\sum_{j=1}^{n} \lambda_j \leq 1$	
NDRS	Add $\sum_{j=1}^{n} \lambda_j \geq 1$	
Efficient Target	$\begin{cases} \hat{x}_{io} = \theta^* x_{io} - s_i^{-*} & i = 1,2,...,m \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r = 1,2,...,s \end{cases}$	$\begin{cases} \hat{x}_{io} = x_{io} - s_i^{-*} & i = 1,2,...,m \\ \hat{y}_{ro} = \phi^* y_{ro} + s_r^{+*} & r = 1,2,...,s \end{cases}$

The dual programs to the envelopment models in Table 5-1 are called multiplier models. Table 5-2 summarizes the multiplier models with respect to the orientations and frontier types.

Table 5-2. Multiplier Models

Frontier Type	Input-Oriented	Output-Oriented
	$\max \sum_{r=1}^{s} \mu_r y_{ro} + \mu$ subject to $\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} + \mu \leq 0$ $\sum_{i=1}^{m} \nu_i x_{io} = 1$ $\mu_r, \nu_i \geq 0(\varepsilon)$	$\min \sum_{i=1}^{m} \nu_i x_{io} + \nu$ subject to $\sum_{i=1}^{m} \nu_i x_{ij} - \sum_{r=1}^{s} \mu_r y_{rj} + \nu \geq 0$ $\sum_{r=1}^{s} \mu_r y_{ro} = 1$ $\mu_r, \nu_i \geq 0(\varepsilon)$
CRS	where $\mu = 0$	where $\nu = 0$
VRS	where μ free	where ν free
NIRS	where $\mu \leq 0$	where $\nu \geq 0$
NDRS	where $\mu \geq 0$	where $\nu \leq 0$

5.4. RETURNS TO SCALE ESTIMATION

As we consider additional information available from DEA regarding scale economies, managers are urged to review existing evidence on the existence of scale economies for the service units being evaluated. There are several issues that need to be considered to ensure that the conclusions using VRS are accurate and meaningful.

Caveat 1: There may be economies of scale but not related to outputs being measured, which can result in misinterpretation of the efficiencies and ways to improve efficiency. An example of this was noted in Chapter 2 where there are many hospital studies that attempt to describe the optimal hospital size in terms of number of beds. There are arguments that this is not the appropriate focus, but the focus should be on the optimal number of surgical operations, childbirths, emergency room episodes, and that each of these has separate and more relevant optimal size levels. If one focused on changing the hospital bed size to improve the efficiency of the hospital, the effort would be on the wrong dimensions and it would not really achieve optimal size. While this may be obvious to many in the health care industry, it was not as obvious to those conducting studies of optimal hospital size focusing on number of beds. In DEA, this is best handled by jointly using multiple outputs and multiple inputs with most productive scale size (MPSS) as described on page 55 in the DEA handbook of Cooper, Seiford and Zhu (2004).

Caveat 2: Some of the assertions about use of returns to scale DEA models may be more relevant for large volumes of service units and less precise when dealing with a group of service units attempting to define ways to change their operations to improve efficiency as a means to improve profitability. Variable returns to scale models assume that a small service unit appears less efficient than a large service unit because there economies of scale; or a large service unit appears less efficient than a small unit because there are diseconomies of scale. This overlooks a real possibility that regardless of whether there are variable or constant returns to scale in the service being evaluated, the small unit may appear more efficient than the large unit because the large unit is operating inefficiently due to poor management, poor method, and weaknesses unrelated to scale. This could be true for the small unit that appears inefficient compared to the large unit, where the small unit may just be over using the resources. In banking for example, there are many studies that suggest that there are no economies of scale. If that is the case and we apply VRS models, operating efficiencies could be hidden. Several

DEA applications have identified cases where large service units were found to be inefficient compared to smaller ones where the large unit could improve by adopting some of the methods used by the smaller units and visa versa. This caveat suggests that when using a VRS model, that the results be compared with and analyzed in conjunction with the basic CRS model.

As demonstrated in Figure 5-1, the VRS envelopment model identifies the VRS frontier with service units exhibiting IRS, CRS, and DRS, respectively. In fact, the economic concept of RTS has been widely studied within the framework of DEA. RTS have typically been defined only for single output situations. DEA generalizes the notion of RTS to the multiple-output case.

It is meaningful to discuss RTS for DMUs (decision making units, or service units) located on the VRS frontier. We discuss the RTS for non-frontier DMUs by their VRS efficient targets as indicated in Table 5-1. Because a VRS model can be either input-oriented or output-oriented, we may obtain different efficient targets and RTS classifications for a specific non-frontier DMU.

Suppose we have five DMUs, A, B, C, D, and H as shown in Figure 5-4. Ray OBC is the CRS frontier. AB, BC and CD constitute the VRS frontier, and exhibit IRS, CRS and DRS, respectively. B and C exhibit CRS. On the line segment AB, IRS prevail to the left of B. On the line segment CD, DRS prevail to the right of C.

Consider non-frontier DMU H. If the input-oriented VRS envelopment model is used, then H′ is the efficient target, and the RTS classification for H is IRS. If the output-oriented VRS envelopment model is used, then H″is the efficient target, and the RTS classification for H is DRS.

However some IRS, CRS and DRS regions are uniquely determined no matter which VRS model is employed. They are region 'I' – IRS, region 'II' – CRS, and region 'III' – DRS. In fact, we have six RTS regions as shown in Figure 5-5. Two RTS classifications will be assigned into the remaining regions IV, V and VI. Region 'IV' is of IRS (input-oriented) and of CRS (output-oriented). Region 'V' is of CRS (input-oriented) and of DRS (output-oriented). Region 'VI' is of IRS (input-oriented) and of DRS (output-oriented).

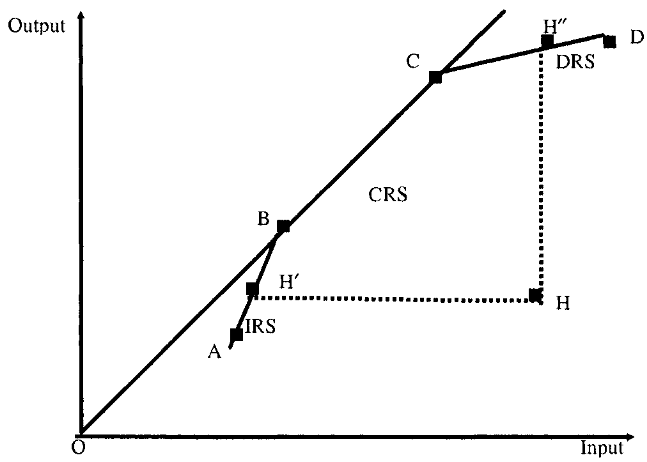


Figure 5-4. RTS and VRS Efficient Target

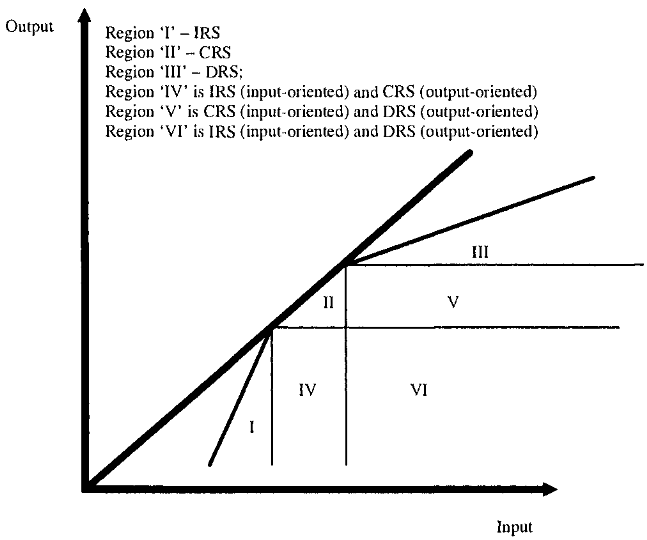


Figure 5-5. RTS Region

As discussed in Seiford and Zhu (1999a), there a number of DEA-based RTS methods. We here present the simple approach developed by Zhu and Shen (1995) and Seiford and Zhu (1999a) and then demonstrate how *DEAFrontier* performs the RTS classification.

Let λ_j^* be the optimal values in CRS envelopment models. We have (i) If $\sum_j^n \lambda_j^* = 1$ in *any* alternate optima, then CRS prevail on DMU_o; (ii) If $\sum_j^n \lambda_j^* < 1$ for *all* alternate optima, then IRS prevail on DMU_o; and (iii) If $\sum_j^n \lambda_j^* > 1$ for *all* alternate optima, then DRS prevail on DMU_o.

Table 5-3. RTS Example

DMU	input 1 (x1)	input 2 (x2)	output (y)	RTS
1	2	5	2	CRS
2	2	2	1	CRS
3	4	1	1	CRS
4	2	1	1/2	IRS
5	6	5	5/2	DRS

However, because of the possible multiple optimal λ_j^* in the CRS envelopment models, it is very difficult to apply the above criteria. Because, in real world applications, the examination of alternative optima is a laborious task, and one may attempt to use a single set of resulting optimal solutions in the application of the RTS methods. Unfortunately, this may yield erroneous results. For instance, consider the example in Table 5-3 taken from Zhu (2003). Table 5-4 reports the optimal solutions obtained from the CRS envelopment model.

Table 5-4. RTS Solutions

DMU	λ_j^*
1	$\lambda_1^* = 1$; $\sum_{j=1}^6 \lambda_j^* = 1$
2	solution 1: $\lambda_2^* = 1$; $\sum_{j=1}^6 \lambda_j^* = 1$
	solution 2: $\lambda_1^* = 1/3$, $\lambda_3^* = 1/3$; $\sum_{j=1}^6 \lambda_j^* = 2/3$
3	$\lambda_3^* = 1$; $\sum_{j=1}^6 \lambda_j^* = 1$
4	$0 \le \lambda_1^* \le 1/12$, $\lambda_2^* = 1/4 - 3\lambda_1^*$, $\lambda_3^* = 1/4 + \lambda_1^*$
	$5/12 \le \sum_{j=1}^6 \lambda_j^* \le 1/2$
5	$\lambda_1^* = 35/48 - \lambda_2^*/3$, $0 \le \lambda_2^* \le 35/16$, $\lambda_3^* = 25/24 - \lambda_2^*/3$
	$85/48 \le \sum_{j=1}^6 \lambda_j^* \le 15/6$

If we obtain $\lambda_1^* = \lambda_3^* = 1/3$ for DMU2, then DMU2 may erroneously be classified as having IRS because $\sum \lambda_j^* < 1$ in one particular alternate solution.

Seiford and Zhu (1999a) show that (i) If DMU_o exhibits IRS, then $\sum_j^n \lambda_j^* < 1$ for *all* alternate optima, and (ii) If DMU_o exhibits DRS, then $\sum_j^n \lambda_j^* > 1$ for *all* alternate optima.

The significance of this finding lies in the fact that the possible alternate optimal λ_j^* obtained from the CRS envelopment models only affect the estimation of RTS for those DMUs that truly exhibit CRS, and have nothing to do with the RTS estimation on those DMUs that truly exhibit IRS or DRS. That is, if a DMU exhibits IRS (or DRS), then $\sum_j^n \lambda_j^*$ must be less (or greater) than one, no matter whether there exist alternate optima of λ_j.

Further, we can have a very simple approach to eliminate the need for examining all alternate optima.

- The CRS efficiency score is equal to the VRS efficiency score *if and only if* CRS prevail on DMU_o. Otherwise,
- $\sum_j^n \lambda_j^* < 1$ *if and only if* IRS prevail on DMU_o.
- $\sum_j^n \lambda_j^* > 1$ *if and only if* DRS prevail on DMU_o.

Thus, in empirical applications, we can explore RTS in two steps. First, select all the DMUs that have the same CRS and VRS efficiency scores regardless of the value of $\sum_j^n \lambda_j^*$. These DMUs are in the CRS region. Next, use the value of $\sum_j^n \lambda_j^*$ (in any CRS envelopment model outcome) to determine the RTS for the remaining DMUs. We observe that in this process we can safely ignore possible multiple optimal solutions of λ_j.

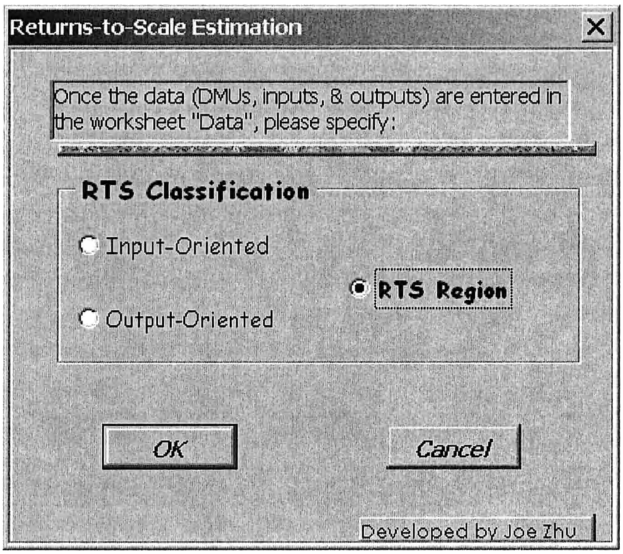


Figure 5-6. RTS Estimation

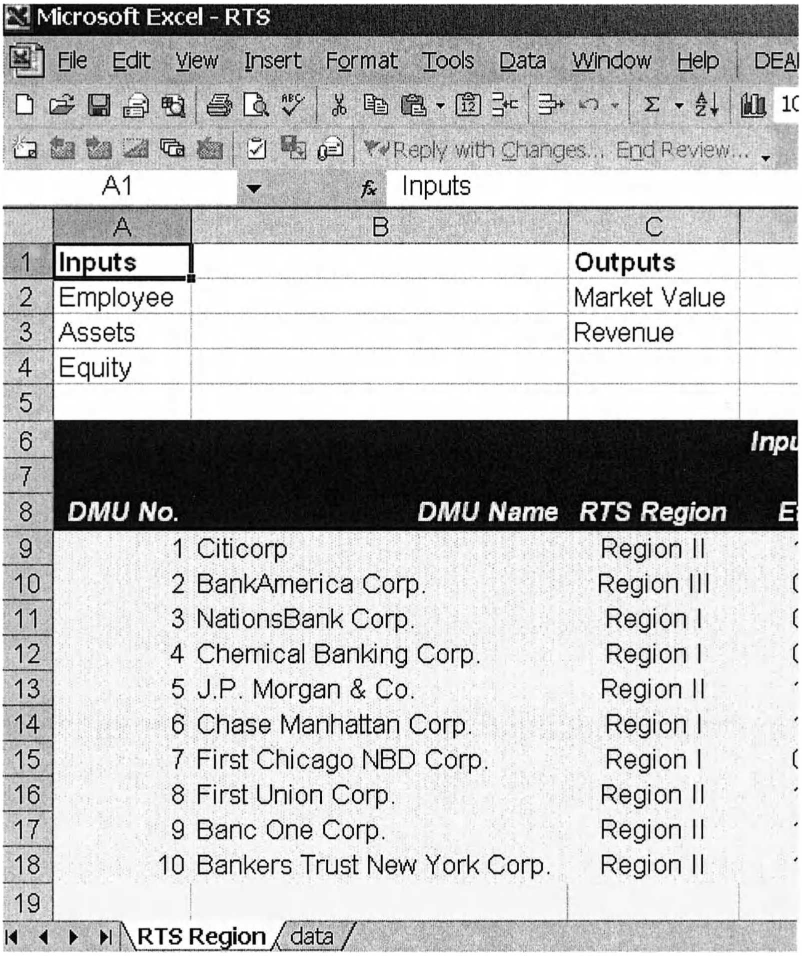


Figure 5-7. RTS Results

To perform the RTS estimation using the *DEAFrontier*, select the "RTS Estimation" menu item and you will be prompted to specify a choice as shown in Figure 5-6. Select "Input-Oriented" if you want to have a RTS classification based upon input-oriented DEA models. Select "Output-Oriented" if you want to have a RTS classification based upon output-oriented DEA models. If "RTS Region: is selected, then the program yields both the RTS regions and the CRS and VRS efficiency scores as shown in Figure 5-7 (see file "RTS.xls" for detailed results.

5.5. RESTRICTED MULTIPLIERS

In the basic multiplier DEA model, the only restriction on the multipliers applied to inputs and outputs, u and v, is the positivity of the multipliers. (This is often imposed by a constraint that the multipliers are greater than or equal to ε, an infinitesimal non-zero number.) Requiring the multipliers to be non-zero is significant from a management perspective, as it requires that all inputs and outputs of the unit be assigned some value. If these weights could be zero, then units that overuse certain inputs would simply assign as zero weight to those to generate a higher performance efficiency rating. It is tantamount to allowing performance to be measure based only on the things the service unit does well, and ignoring the things it does poorly.

With only the constraint that the multipliers be greater than zero, the resulting weights, which are effectively relative prices, may not reflect known relationships in relative prices or values. Management may not know the relative value of price of every output and input, but they may have some insight about a ranking of value and even some specific ideas of relative value.

For example, a hospital surgical unit may do appendectomies and cardiac bypass surgery, but they generally do not have an efficient standard indicating the relative amount of resources needed to remove an appendix versus heart bypass operations. However, almost anyone with basic understanding of these two operations would agree that bypass surgery demands more resources. This insight would allow us to add the constraint that the weight on the bypass surgery should be greater than the weight for appendectomy operations completed. Adding this constraint, beyond the basic constraint that the weights be non-zero, causes the DEA analysis to value the bypass at least as highly as appendectomy. This increases the ability of DEA to identify finer distinctions in efficiency than the basic model and makes the results more powerful in that the inefficiencies that can be remedied may increase with the added constraints. If the medical staff agrees that the bypass surgery should be valued a 2, 3, or Q times greater than appendectomies, the constraint that bypass have a higher value can be replaced with the constraint that the value of bypasses be equal to or greater than 2, 3 or Q times the value of appendectomies.

Adding these weights (as constraints or restrictions) would further increase the power of the model and offer more options and paths to improve service unit performance. Even in cases where the relative prices may be determinable, it has been useful to include relative weights. For example, full time equivalents (FTEs) for different types of personnel may be included as resources used by a service unit. The pay scale for units in different parts of the world may differ and even for similar grade personnel, the pay will differ by seniority, length of service as well as ability. Rather than try to develop weights based on individual salary levels, constraints are added to reflect general differences in salary levels. For instance, in a nursing home, the constraint that the manager salary is equal to or greater than Z times the registered nurses might be included, forcing the weight on the manager to be at least Z times the weight of the registered nurse input. Similar constraints are added for any relations that are known and reliable such as the relative cost of food preparation staff, cleaning staff, nurse's aids etc.

By adding the known information about relative values, we can add constraints to the DEA formulation that will increase its power to identify inefficiencies and may even identify additional units that are inefficient. The result is that this can increase the insights available from DEA and these increases can dramatically increase the value of the insights generated by DEA. These constraints have also been incorporated in different ways and have been generated from different sources. Some have tried to develop the range of acceptable weights based on management judgment and expertise while others try to infer relative weights from the data. In short, the manger should incorporate only those weight constraints that they have carefully evaluated and believe reflect real and reliable differences in the relative values of the inputs and outputs.. Otherwise, the service unit that is identified as inefficient could challenge these weights and, if the constraints are found to be questionable, the conclusions of the DEA analysis can be discounted or deemed useless. These constraints are sometimes referred to as weight restrictions, assurance regions, cone ratios. While these terms can be associated with somewhat different ways of incorporating these additional constraints in the DEA model, they are all methods to set boundaries on the weights and incorporate these as a constraint in the DEA model.

In the DEA literature, a number of approaches have been proposed to introduce additional restrictions on the values that the multipliers can assume.

We here present the assurance region (AR) approach of Thompson et al. (1990). To illustrate the AR approach, suppose we wish to incorporate additional inequality constraints of the following form into the multiplier DEA models as given in Table 5-2:

$$\alpha_i \le \frac{v_i}{v_{i_o}} \le \beta_i, \qquad i = 1,\dots,m$$

$$\delta_r \le \frac{\mu_r}{\mu_{r_o}} \le \gamma_r, \qquad r = 1,\dots,s$$

Here, v_{i_o} and μ_{r_o} represent multipliers which serve as "numeraires" in establishing the upper and lower bounds represented here by α_i, β_i, and by δ_r, γ_r for the multipliers associated with inputs i =1, …, m and outputs r = 1,, s where α_{i_o} = β_{i_o} = δ_{r_o} = γ_{r_o} = 1. The above constraints are called Assurance Region (AR) constraints as in Thompson et al. (1990).

Uses of such bounds are not restricted to prices. For example, Zhu (1996a) uses an assurance region approach to establish bounds on the weights obtained from uses of Analytic Hierarchy Processes in Chinese textile manufacturing in order to reflect how the local government is measuring the textile manufacturing performance.

The generality of these AR constraints provides flexibility in use. Prices, units and other measures may be accommodated and so can mixtures of such concepts. Moreover, one can first examine provisional solutions and then tighten or loosen the bounds until one or more solutions are attained that appears to be reasonably satisfactory to decision makers who cannot state the values for their preferences in an a priori manner.

	A	B	C	D	E	F	G
1	Banks	Employee	Assets	Equity		Market Value	Revenue
2	Citicorp	85300	256853	19581		33221.7	31690
3	BankAmerica Corp.	95288	232446	20222		27148.6	20386
4	NationsBank Corp.	58322	187298	12801		20295.9	16298
5	Chemical Banking Corp.	39078	182926	11912		16971.3	14884
6	J.P. Morgan & Co.	15600	184879	10451		15003.5	13838
7	Chase Manhattan Corp.	33365	121173	9134		12616.4	11336
8	First Chicago NBD Corp.	35328	122002	8450		12351.1	10681
9	First Union Corp.	44536	131880	9043.1		16815	10582.9
10	Banc One Corp.	46900	90454	8197.5		14807.4	8970.9
11	Bankers Trust New York Cor	14000	104000	5000		5252.4	8600
12							

data

Figure 5-8. Bank Example

We next illustrate how to use the *DEAFrontier* to solve the "Restricted Multiplier" models. Consider an example where we have top 10 US commercial banks in 1995 with three inputs (employee, assets and equity) and two outputs (market value and profit), as shown in Figure 5-8. (see Seiford and Zhu (1999b) for detailed discussion on this data set.)

To run the "Restricted Multiplier" models, we need to first set up the sheet "Multiplier" which contains the ARs. For example, if we want to include the following ARs

$$1 \le \frac{v_{Employee}}{v_{Assets}} \le 2.5$$

$$1.5 \le \frac{v_{Employee}}{v_{Equity}} \le 3$$

$$3 \le \frac{\mu_{MarketValue}}{\mu_{\mathrm{Re}venue}} \le 4$$

then the data in the "Multiplier" sheet should be entered as shown in Figure 5-9.

	A	B	C	D	E
1	1	Employee	Assets	2.5	
2	1.5	Employee	Equity	3	
3	3	Market Value	Revenue	4	
4					

data | multiplier

Figure 5-9. Restricted Multipliers

To avoid any errors, we suggest copying and pasting the input and output names from the "data" sheet when entering the information into the "Multiplier" sheet. If the input (output) names in the two sheets do not match, the program will stop.

Once the "Multiplier" sheet is set up, select the "Restricted Multipliers" menu item and you will be prompted to choose a DEA model, as shown in Figure 5-10. Figure 5-7 shows the results of the input-oriented CRS multiplier model with the above ARs. (See also file "Restricted Multiplier.xls" in the attached CD.)

Note that you can also add ARs that link the input and output multipliers for the "Restricted Multipliers". Note also that if the ARs are not properly specified, then the related DEA model may be infeasible. If that happens, the program will return a value "-9999" for the efficiency score.

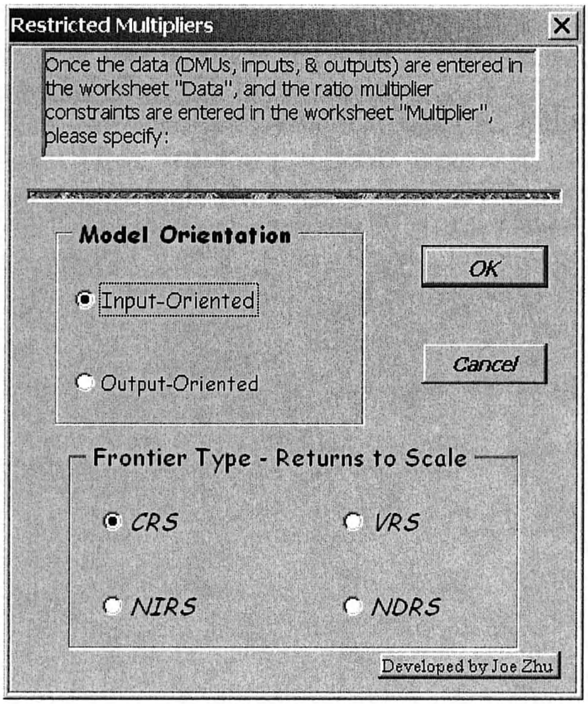
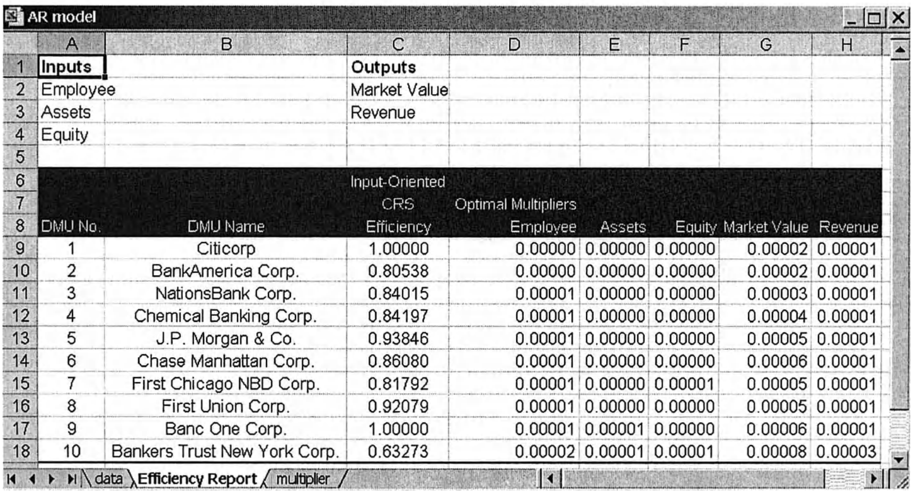


Figure 5-10. Restricted Multiplier Models



AR model

DMU No.	DMU Name	Input-Oriented CRS Efficiency	Optimal Multipliers Employee	Assets	Equity	Market Value	Revenue
1	Citicorp	1.00000	0.00000	0.00000	0.00000	0.00002	0.00001
2	BankAmerica Corp.	0.80538	0.00000	0.00000	0.00000	0.00002	0.00001
3	NationsBank Corp.	0.84015	0.00001	0.00000	0.00000	0.00003	0.00001
4	Chemical Banking Corp.	0.84197	0.00001	0.00000	0.00000	0.00004	0.00001
5	J.P. Morgan & Co.	0.93846	0.00001	0.00000	0.00000	0.00005	0.00001
6	Chase Manhattan Corp.	0.86080	0.00001	0.00000	0.00000	0.00006	0.00001
7	First Chicago NBD Corp.	0.81792	0.00001	0.00000	0.00001	0.00005	0.00001
8	First Union Corp.	0.92079	0.00001	0.00000	0.00000	0.00005	0.00001
9	Banc One Corp.	1.00000	0.00001	0.00001	0.00000	0.00006	0.00001
10	Bankers Trust New York Corp.	0.63273	0.00002	0.00001	0.00001	0.00008	0.00003

Inputs: Employee, Assets, Equity. Outputs: Market Value, Revenue.

Figure 5-11. Restricted Multiplier Model Result

Including weight restrictions noticeably enhances the insights provided by DEA, and we highly recommend including these constraints where at least some of the relationships are known, clear, and widely accepted as reasonable. However, we also suggest that the results with constraints on the weights be compared with the results of the basic DEA model. The method and benefit of adding weight constraints is illustrated chapter 9, which describes the application of DEA to the Government of Canada purchasing services.

5.6. MEASURE-SPECIFIC MODELS

Although DEA does not need *a priori* information on the underlying functional forms and weights among various input and output measures, it assumes proportional improvements of inputs or outputs. This assumption becomes invalid when a preference structure over the improvement of different inputs (outputs) is present in evaluating (inefficient) DMUs (see Zhu, 1996). We need models where a particular set of performance measures is given pre-emptive priority to improve.

Table 5-5. Measure-specific Models

Frontier Type	Input-Oriented	Output-Oriented
CRS	$\min \theta - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io} \quad i \in I;$ $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i \notin I;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1,2,\ldots,s;$ $\lambda_j \geq 0 \quad j = 1,2,\ldots,n.$	$\max \phi - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1,2,\ldots,m;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro} \quad r \in O;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r \notin O;$ $\lambda_j \geq 0 \quad j = 1,2,\ldots,n.$
VRS	Add $\sum_{j=1}^{n} \lambda_j = 1$	
NIRS	Add $\sum_{j=1}^{n} \lambda_j \leq 1$	
NDRS	Add $\sum_{j=1}^{n} \lambda_j \geq 1$	
Efficient Target	$\begin{cases} \hat{x}_{io} = \theta^* x_{io} - s_i^{-*} & i \in I \\ \hat{x}_{io} = x_{io} - s_i^{-*} & i \notin I \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r = 1,2,\ldots,s \end{cases}$	$\begin{cases} \hat{x}_{io} = x_{io} - s_i^{-*} & i = 1,2,\ldots,m \\ \hat{y}_{ro} = \phi^* y_{ro} + s_r^{+*} & r \in O \\ \hat{y}_{ro} = y_{ro} + s_r^{+*} & r \notin O \end{cases}$

Let $I \subseteq \{1,2, \ldots, m\}$ and $O \subseteq \{1,2, \ldots,s\}$ represent the sets of specific inputs and outputs of interest, respectively. Based upon the envelopment models, we can obtain a set of measure-specific models where only the inputs associated with I or the outputs associated with O are optimized (see Table 5-5).

The measure-specific models can be used to model uncontrollable or non-discretionary inputs and outputs (see Banker and Morey (1986)). The controllable measures are related to set I or set O.

A DMU is efficient under envelopment models if and only if it is efficient under measure-specific models, i.e., both the measure-specific models and the envelopment models yield the same frontier. However, for inefficient DMUs, envelopment and measure-specific models yield different efficient targets.

Consider Figure 2-1 in chapter 2. If the teller hours input is of interest (i.e., if the management is only interested in reducing the teller hours), then the measure-specific model will yield the efficient target of B4 for inefficient B2. If the supply dollars input is of interest, B4 will be the target for B1. The envelopment model (2.2) projects B1 or B2 to the frontier by reducing the two inputs proportionally.

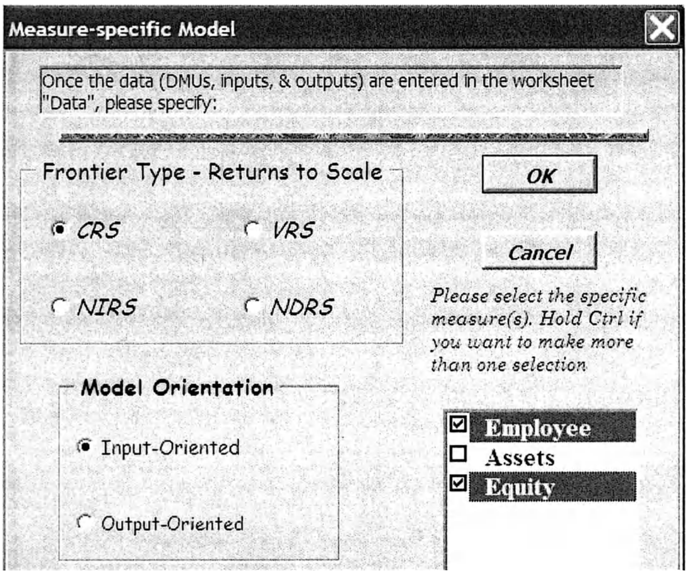


Figure 5-12. Measure Specific Models

To run the measure-specific models using DEAFrontier, select the "Measure Specific Model" menu item. It prompts a model selection window as shown in Figure 5-12. Figure 5-12 uses the bank example in Figure 5-8. If we select Employee and Equity, this means we are only interested in these two inputs and the resulting efficiency scores only reflect the possible reduction in these two inputs.

5.7. SLACK-BASED MODELS

The input-oriented DEA models consider the possible (proportional) input reductions while maintaining the current levels of outputs. The output-oriented DEA models consider the possible (proportional) output augmentations while keeping the current levels of inputs. Charnes, Cooper, Golany, Seiford and Stutz (1985) develop an additive DEA model which considers possible input decreases as well as output increases simultaneously. The additive model is based upon input and output slacks. For example,

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

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \qquad i = 1,2,\ldots,m; \tag{5.1}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \qquad r = 1,2,\ldots,s;$$

$$\lambda_j, s_i^-, s_r^+ \geq 0$$

Note that model (5.1) assumes equal marginal worth for the nonzero input and output slacks. Therefore, caution should be excised in selecting the units for different input and output measures. Some *a priori* information may be required to prevent an inappropriate summation of non-commensurable measures. Management's experience and expert opinion, critical in all productivity analysis, may be used to assess the relative importance and value of identifying input and output slacks. (see Seiford and Zhu (1998)).

Model (1.8) therefore is modified to a weighted CRS slack-based model as follows (Ali, Lerme and Seiford, 1995; Thrall, 1996).

$$\max \sum_{i=1}^{m} w_i^- s_i^- + \sum_{r=1}^{s} w_r^+ s_r^+$$

subject to (5.2)

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \qquad i = 1,2,\ldots,m;$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \qquad r = 1,2,\ldots,s;$$

$$\lambda_j, s_i^-, s_r^+ \geq 0$$

where w_i^- and w_r^+ are user-specified weights obtained through value judgment. The DMU_o under evaluation will be termed efficient *if and only if* the optimal value to (5.2) is equal to zero. Otherwise, the nonzero optimal s_i^{-*} identifies an excess utilization of the *i*th input, and the non-zero optimal s_r^{+*} identifies a deficit in the *r*th output. Thus, the solution of (5.2) yields the information on possible adjustments to individual outputs and inputs of each DMU. Obviously, model (5.2) is useful for setting targets for inefficient DMUs with *a priori* information on the adjustments of outputs and inputs.

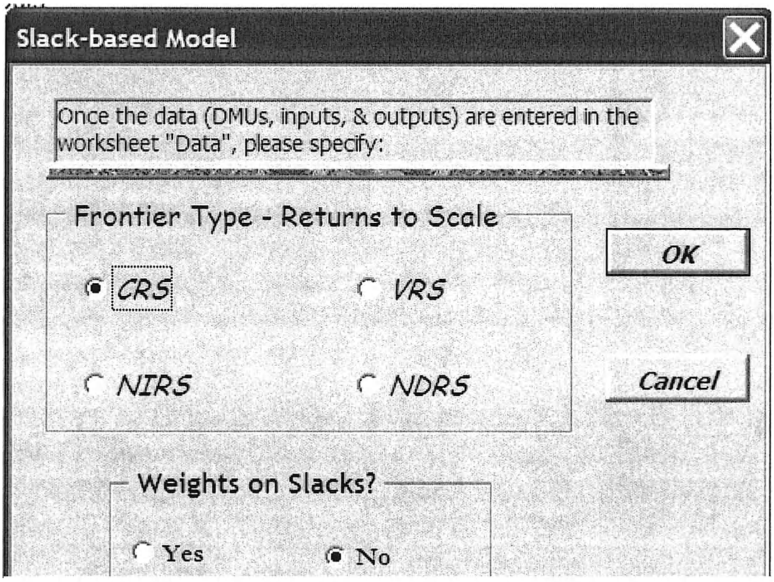
Table 5-6. Slack-based Models

Frontier type	Slack-based DEA Models
CRS	$\max \sum_{i=1}^{m} w_i^- s_i^- + \sum_{r=1}^{s} w_r^+ s_r^+$ subject to $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1,2,\ldots,m;$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1,2,\ldots,s;$ $\lambda_j, s_i^-, s_r^+ \geq 0$
VRS	Add $\sum_{j=1}^{n} \lambda_j = 1$
NIRS	Add $\sum_{j=1}^{n} \lambda_j \leq 1$
NDRS	Add $\sum_{j=1}^{n} \lambda_j \geq 1$

One should note that model (5.2) does not necessarily yield results that are different from those obtained from the model (5.1). In particular, it will not change the classification from efficient to inefficient (or vice versa) for any DMU.

Model (5.2) identifies a CRS frontier, and therefore is called CRS slack-based model. Table 5-6 summarizes the slack-based models in terms of the frontier types

To run the measure-specific models using DEAFrontier, select the "Slack-based Model" menu item. It prompts a model selection window as shown in Figure 5-13. If "Yes" is selected for the specifying the slack weights, then it prompts the user to specify the weights as shown in Figure 5-13.



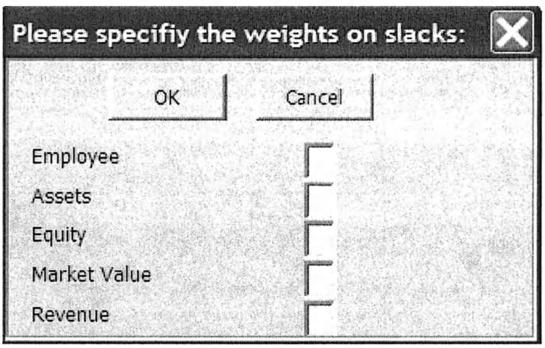


Figure 5-14. Slack Weights

5.8. OTHER DEA MODELS

The above introduces some DEA extensions. However, there are also other DEA models and approaches that can be useful to service activities. The following briefly discusses some other DEA models and approaches. Interested readers can lean more about these models and approaches in the cited references and later chapters. An Excel Solver based DEA software similar to the DEAFrontier provided in the current book is available at Zhu (2003).

Färe, Grosskopf and Lovell (1994) develop DEA-based Malmquist productivity models where one can measure the productivity changes over time. The productivity change can be decomposed into a frontier shift component and technical change component. The Malmquist productivity model requires (i) at least two time periods data, (ii) the DMUs over time have to remain the same, and (iii) the number of DMUs in each period must be the same.

Super-efficiency models where a DMU under evaluation is removed from the reference set were developed for the purpose of ranking efficient units (Andersen and Petersen, 1993). For the input-oriented CRS model, the super-efficiency scores for efficient units will be greater than one. The scores for inefficient units will not change from the CRS model to the super-efficiency model. However, cautions should be used when we use the super-efficiency model. As noted in Seiford and Zhu (1999c), some super-efficiency models can be infeasible. Also, super-efficiency may over-characterize a unit's performance (see Chen and Sherman (2004), Chen (2004, 2005).)

Seiford and Zhu (1999b, 2003) and Zhu (2003) introduce a new DEA approach called context-dependent DEA. Adding or deleting an inefficient DMU or a set of inefficient DMUs does not alter the efficiencies of the existing DMUs and the best-practice frontier. The inefficiency scores change only if the best-practice frontier is altered. i.e., the performance of DMUs depends only on the identified best-practice frontier. In contrast, researchers of the consumer choice theory point out that consumer choice is often influenced by the context. e.g., a circle appears large when surrounded by small circles and small when surrounded by larger ones. Similarly a product may appear attractive against a background of less attractive alternatives and unattractive when compared to more attractive alternatives (Simonson and Tversky, 1992). The context-dependent DEA considers and models this influence within the framework of DEA. Relative attractiveness of a unit depends on the evaluation context constructed from alternative options (or DMUs). In the context-dependent DEA, a set of DMUs is grouped into different levels of best-practice frontiers. Each best-practice frontier provides an evaluation context for measuring the relative attractiveness. e.g., the second-level best-practice frontier serves as the evaluation context for measuring the relative attractiveness of the DMUs located on the first-level (original) best-practice frontier. It can be seen that the presence or absence (or the shape) of the second-level best-practice frontier affects the relative attractiveness of DMUs on the first-level best-practice frontier. A relative attractiveness measure is obtained when DMUs having worse performance are chosen as the evaluation context. When DMUs in a specific level are viewed as having equal performance, the attractiveness measure allows us to differentiate the "equal performance" based upon the same specific evaluation context (or third option). The context-dependent DEA can rank the units in different levels of best-practice frontiers and is also useful in step-wise performance improvement.

Zhu (2003) develops a number of new DEA models that can be used in benchmarking a set of units against a given set of (efficient or standard) units. The DEA benchmarking models developed in Zhu (2003) enables the user to compare the performance of two sets of DMUs and compare each DMU to the same standard or benchmark.

Quality Adjusted DEA (Q-DEA) – Incorporating the quality dimension in DEA productivity analysis.

Data envelopment analysis (DEA) has been demonstrated to be a powerful benchmarking methodology for organizations in which multiple inputs and outputs need to be assessed to identify best practices and improve productivity. Most DEA benchmarking studies have excluded quality, even in service-sector applications such as health care where quality is a key element of performance. One will note that in all the discussion of service applications above, no mention of quality of service was considered. This limits the practical value of DEA in organizations where maintaining and improving service quality is critical to achieving performance objectives. Indeed, sectors like health care often are thought to have quality as the key focus with only secondary concern for productivity. As a patient, we all are likely to hope that quality is more important than efficiency when we visit a physician or are admitted to a hospital. All the applications of DEA that ignore quality implicitly assume that 1) quality is unimportant, 2) all units being compared have all achieved a high quality status and that dimension needs not attention, 3) quality is totally independent of productivity, or 4) no quality measures exist making that type of analysis unfeasible.

Alternative methods incorporating quality in DEA benchmarking are demonstrated and evaluated in some detail in Chapter 7. It is shown that simply treating the quality measures as DEA outputs does not help in discriminating the performance. In chapter 7 we describe what we believe it the most sensitive method of including quality in DEA benchmarking: quality-adjusted DEA (Q-DEA) as of the time of publication of this book.

Chapter 7 reports the results of applying Q-DEA to a U.S. bank's 200-branch network that required a method for benchmarking to help manage operating costs and service quality. Q-DEA findings helped the bank achieve cost savings and improved operations while preserving service quality, a dimension critical to its mission. New insights about ways to improve branch operations based on the best-practice (high-quality low-cost) benchmarks identified with Q-DEA are also described. This demonstrates the practical need and potential benefits of Q-DEA and its efficacy in one application, and underscores the need for further research on measuring and incorporating quality into DEA benchmarking.

Data Sensitivity Analysis

DEA is a deterministic approach and can be sensitive to the errors in the data. Most of the discussion in this book assumes that the inputs and outputs will reflect all those critical to the service process and that the data specified would be collected will be accurate. We have cautioned that misspecification or an incomplete model can make the analysis misleading and lead to poor decision that can be costly.

Using inaccurate data with DEA, as with any quantitative technique, can generate misleading results and, if these errors are not recognized, the results could generate dysfunctional decisions. In this volume, we assume that the data quality would be carefully scrutinized before using it for the DEA analysis. However, the impact of data errors would not be identical to the impact of same errors using other techniques. For example, DEA is very dependent on having reliable data for the ERS units, as they are the basis for measuring efficiency scores.

There are a number of DEA sensitivity approaches available to perform various sensitivity analyses and gain insights about whether idiosyncrasies in the data may skew the results due to unintended data misspecifications. The interested reader is referred to Cooper, Li, Seiford and Zhu (2004).

While we have not described every alternative DEA model, and while other variations may be developed to adapt to research or different types of applications, the models covered are the ones we believe are most relevant for managers based on presentations and research published through 2006.

Chapter 6

MANAGING BANK PRODUCTIVITY

This chapter describes a bank that used DEA to make real and substantial improvement in its branch productivity, and profits while maintaining service quality. Over $6 Million of annual expense savings not identifiable with traditional financial and operating ratio analysis were identified in this 33-branch system.

DEA explicitly considers all the resources each branch uses and the services it provides. It compares branches objectively to identify the best practice branches, the less productive branches, and the changes the less productive branches need to implement to reach the best practice level and improve their profitability.

6.1. INTRODUCTION

A reorganization of an 80 branch banking system resulted in a 30 percent reduction of personnel with no reduction in service quality (Sherman (1989)). By streamlining its branch operations, a brokerage firm freed up more than 20 percent of its annual operating costs and used these savings to expand marketing and new branch operations for future business development (Bank Technology Report 1992). Following a series of acquisitions, a 350-branch tri-state commercial bank achieved similar results (Iida (1991)). In each case, management employed DEA to develop a path to improve productivity.

It is a particularly difficult task for a service business to improve its operations effectively and find substantial cost savings without sacrificing quality of service. Unlike a manufacturing concern, a service business has a number of subjective factors that affect its productivity and service quality. In a bank, this includes customers' needs and behavior in receiving the service, service provider's judgment and skill in providing the service, and the changing mix of services provided. In the case of a northeastern bank, which we will call Growth Bank, DEA (Data Envelopment Analysis) highlighted many areas for improvement.

Growth Bank was experiencing an extended period of sustained growth reflecting its successful efforts to adapt and market its financial services in a relatively competitive market. The managers wanted to expand its branch system beyond the existing 33 branches to cover a wider geographic region. They wanted to streamline existing operations without sacrificing service quality to ensure that the costly resources for branch operations were utilized efficiently. They hoped to obtain some of the resources they needed to expand the system by saving costs for the existing network. Professor Sherman was consulted to determine whether applying DEA could help identify ways to improve branch productivity and profits beyond the level achieved by Growth Bank's personnel.

The managers were firmly committed to focusing on productivity gains to improve profitability, which had been declining as lending operations weakened. Branch profit and loss statements provided limited insight into the operating productivity of each branch. Some small urban branches generated disproportionately high profits, reflecting clusters of large transactions by a few unrepresentative large corporate customers. Management did not know how to locate the more efficient branches. In fact, in our analysis we found that several of these highly profitable branches had substantial operating inefficiencies. Branches that primarily provided check-cashing functions reported relatively low profitability but were not necessarily inefficient. In addition to reducing costs of existing branches, management wanted to develop models of well-run branches to plan more effectively for new offices.

6.2. APPLYING DEA TO GROWTH BANK

In our review, we explicitly considered all the resources used to support branch activity. These included three types of personnel (the dominant resource used by branches); office space; and other direct branch operating costs, such as supplies, utilities, and marketing expenses. Management sought to determine the ideal staffing level, office layout, and support cost level for each branch based on the volume and mix of services it intended to offer. Staffing varied among branches, but all had (1) sales-service, platform staff responsible for more complex transactions, loan services, and new accounts, (2) one branch manager, and (3) customer service, teller personnel responsible for high volume, less complex transactions such as deposits, bank checks, bond transactions, and traveler's checks.

At the time, management was using ratios, such as transactions per teller, cost per transaction, and branch size measured in terms of deposits, to address productivity and establish staffing levels. While these measures are commonly used in banking, they did not help management to be sensitive to the type and mix of transactions nor to consider more complex non teller transactions such as new accounts, loan applications and ATM servicing.

The strength of data envelopment analysis is its ability to consider explicitly use of multiple resources used to provide multiple services while comparing branches. It is a highly objective benchmarking technique particularly well suited to such multi-office service organizations as bank branches operations, as Sherman and Gold (1985) first demonstrated. More general applicability to bank performance has also been demonstrated by Berg, Førsund, and Jansen (1990), Charnes et al. (1990), and Yue (1992). Unlike many benchmarking approaches that rely on managers to observe, compare and identify best practice techniques, DEA helps the user to identify best practices that are too complex to be identified through observation and traditional analytic techniques. It enables management to determine objectively the best practices in complex service operations. The best practice service provider is the one that uses the least resources to provide its volume and mix of service at or above the quality standard of the business. Service costs decline as the less productive service operations are improved to the best practice level, guided by DEA.

Many managers of service organizations would describe benchmarking and best practice analysis as basic, widely accepted concepts already used in their businesses. Closer examination indicates that the traditional techniques used to identify and promulgate best practices are not very effective, largely because their operations are too complex to allow them to identify best practices accurately. DEA provides an objective way to identify best practices in these service organizations and has consistently generated new insights that lead to substantial productivity gains that were not otherwise identifiable.

While traditional approaches successfully identify ways to marginally improve operations throughout a business, this approach seeks to achieve a quantum leap in productivity of the service delivery system.

While some organizations may employ DEA primarily to boost short-term profitability, Growth Bank's objective was to redeploy cost savings to expand the service distribution system and to improve customer service to enhance market share and long-term profitability.

DEA compares the actual operating results of each service unit with those of all other service units and identifies the less productive units - those that are operating inefficiently. A less productive service unit is defined as one that can produce its volume, mix, and quality of outputs with fewer resources based on comparison with the best practice units. It identifies the best practice units (those that are not less efficient than other units being evaluated) and measures the magnitude of inefficiency of the less productive units compared to the best practice units. DEA uniquely obtains these insights by explicitly considering the volume and mix of resources used and the volume and mix of services provided by each service unit. The best practice units are relatively efficient and are identified by a DEA productivity rating of 100 percent ($\theta = 1$). The inefficient less productive units are identified by a productivity rating of less than 100 percent ($\theta<1$).

Specifically, DEA determines the following:

- the best practice - most productive group of service units;
- the less productive service units compared to the best practice units;
- the amount of excess resources used by each of the less productive units;
- the amount of excess capacity or ability to increase service outputs in less productive units without utilizing added resources; and
- the set of best practice service units most similar to the less productive units. This peer group, referred to as the best practice reference set, most directly indicates that excess resources are being used by the inefficient DMU.

6.3. SPECIFYING RESOURCE INPUTS AND SERVICE OUTPUTS

We used DEA as a lead tool to review productivity because it considers all key services and all resources used by the branch explicitly and objectively. We used five resources and five service transaction groups in this study based on management's' assessment of key branch activities (Figure 6-1).

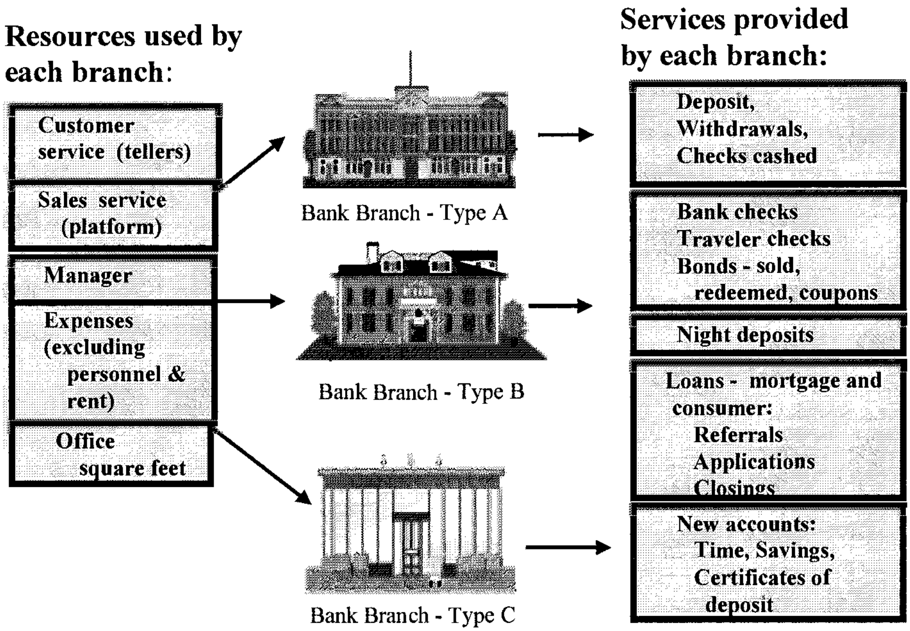


Figure 6-1. Bank Branches

All branch types (A, B and C) use the same set of resources to provide all branch services used for the DEA analysis of Growth Bank's branch productivity. Each branch uses a different amount of each of the resources and offers all of the services. Each branch provides a different volume and mix of these services, depending on their custom demand. Example of branch types include urban, suburban, and shopping mall branches.

Management identified five key resources used by each branch: teller, platform, and manager personnel full time equivalents (FTEs), operating expenses excluding personnel and rent, and square feet of office space. They identified about 20 services provided by each branch. We included 15 of these in the study combined into five sets of services.

The services represented transactions provided by teller and platform personnel and included groups of transaction types considered similar by bank management. For example, we combined deposits, withdrawals, and checks cashed into one group, and treated each deposit, withdrawal, or check cashed as one transaction. We assumed that these transactions are similar enough in terms of resources used on average to be treated as one group of transactions. The DEA productivity rating could be skewed for branches that have deposits consisting of many checks as occurs in branches near retail business centers, because these deposits require more resources than a simple deposit of one or two checks. We excluded other service transactions because they were believed to be of insignificant volume (such as wire transfers) or because the bank did not have adequate records to include the service (such as safe-deposit-box visits).

We collected data on resources used and service volume for the previous year for each branch. Use of a full year eliminated some of the cyclicality issues that might arise with the use of one month or one quarter, though other insights could be gained by analyzing each quarter with DEA.

Before proceeding to use the DEA results, we reviewed the branch productivity ratings with bank management to judge whether excluded services or the transaction groupings might be biasing the productivity assessments of branches. The managers thought this was not a significant issue for this set of branches.

6.4. DEA BRANCH PRODUCTIVITY RESULTS

We use the basic CRS DEA model (2.2) described in chapter 2 with the inputs and outputs described above. The DEA results identified the relatively efficient best practice branches, the less productive branches (Table 6-1), and the magnitude of the excess resources used by less productive branches using the DEA projection (Table 6-2). Branch B1 was less productive with a DEA productivity rating of 66 percent, suggesting that it could provide its current mix and volume of services with only about 66 percent of the resources it actually consumed. Branch B3 had a DEA productivity rating of 37 percent indicating that it was using about 63% excess resources. The analysis showed that Branches such as B2 and B7 were best practice branches, indicated by the DEA productivity rating of 100 percent. In fact, 23 of the 33 branches were using excess resources and 10 were using 30 percent or more excess resources.

These findings indicated that the bank could make substantial productivity improvements and cost reductions. The results also indicated that less productive branches were inefficient compared to particular sets of best practice branches. For example, branch B1 is inefficient specifically in comparison to branches B2, B16, and B19 (referred to as the best practice branch reference set for B1 in Table 6-1). These reference sets indicate the branches that are most similar to the less productive branches in their mix of services and resources, and they are used to identify the specific operating characteristics that make the less productive branch more costly to operate.

Table 6-1. Growth Bank Branch Productivity Ratings

Branch	DEA Productivity Rating	Best Practice Branch (Reference Set)
B1	66%	B2 B16 B19
B2	100%	
B3	37%	B16 B7 B25
B4	46%	B16 B19 B17 B26
B5	92%	B2 B16 B19
B6	48%	B16 B7 B26
B7	100%	
B8	97%	B16 B7 B26
B9	66%	B16 B19 B7 B17 B24
B10	100%	
B11	79%	B16 B19 B17
B12	55%	B16 B7
B13	99%	B2 B16
B14	81%	B16 B7
B15	47%	B16 B7 B25
B16	100%	
B17	100%	
B18	74%	B16
B19	100%	
B20	63%	B16 B19 B17
B21	100%	
B22	79%	B2 B16
B23	67%	B16
B24	100%	
B25	100%	
B26	100%	
B27	73%	B16 B19 B17 B26
B28	66%	B16 B19 B17 B26
B29	97%	B2 B16
B30	77%	B16 B19 B17
B31	73%	B2 B19
B32	86%	B2 B16 B19
B33	71%	B16 B19 B7 B17 B24

For each branch, this table shows the DEA productivity rating based on comparing each branch with every other branch. A rating of 100 percent indicates a best practice branch. A rating of less than 100 percent indicates a less productive branch. The best practice branch reference set for a less productive branch is the set of best practice branches identified with DEA that provide their volume and mix of services with fewer resources than that less productive branch.

Table 6-2. Growth Bank Potential Resource Savings in Less Productive Branches

	Potential Resource Savings[a]				
Less productive branches	Customer Service (tellers)[c]	Sales Service (platform)[c]	Managers[c]	Expenses[b]	Square Feet
B1	4.5	1.8	0.3	$222,928	1,304
B3	1.9	1.6	0.7	$295,989	1,133
B4	2.3	1.4	0.7	$189,745	1,051
B5	0.7	1.9	0.1	$367,020	1,899
B6	1.6	1.4	0.6	$122,474	1,556
B8	1.5	1.3	0.3	$10,526	40
B9	1.2	0.7	0.4	$116,716	976
B11	3.3	0.7	0.2	$329,403	774
B12	1.1	1.1	0.7	$122,433	889
B13	0.1	0.1	0	$131,389	1,477
B14	0.7	2	0.4	$81,024	502
B15	1.1	1.3	0.7	$135,920	961
B18	2	0.8	0.4	$206,693	361
B20	3	1.5	0.4	$280,853	1,176
B22	5.2	1.5	0.3	$496,072	605
B26	1	1	0.3	$112,147	1,491
B27	1.9	3.1	0.3	$188,394	960
B28	1.8	1.6	0	$233,870	60
B29	1.7	0.9	0.2	$176,227	3,669
B30	4	6.1	0.3	$551,272	5,377
B31	0.8	2.7	0.1	$94,692	824
B32	0.9	0.6	0.5	$107,934	480
B33	3.3	10.6	0.2	$2,510,589	2,016
Total savings	45.6	45.7	8.1	$7,084,308	29,580
Total resources used by all branches	203.5	144.0	32.5	$20,997,872	103,508
Savings as % of total resources	22.4%	31.7%	25.1%	33.7%	28.6%

(a) Potential resource reductions, which would make less productive branches as efficient as the best practice branches.

(b) Operating expenses excluding personnel and occupancy expenses.

(c) Full time equivalent personnel savings

The DEA results (based upon DEA projection) in Table 6-3 indicate that the less productive branches could increase their service transaction volume in addition to the resource savings reported in Table 6-2, if they became as efficient as the best practice branches. The actual transaction volume for each less productive branch is reported. The potential service volume expansion indicates the amount of added transactions these branches could handle in addition to reducing their resources by the amounts reported in Table 6-2.

Table 6-3. Growth Bank Service Volume and Potential Volume Increase in Addition to Potential Resource Savings

	ACTUAL SERVICE TRANSACTION VOLUME					POTENTIAL SERVICE VOLUME EXPANSION				
Less Productive Branch	Deposits Withdrawals Checks Cashed	Bank Checks Bonds Traveler Checks	Loans: Applications Closings	Night Deposits	New Accounts	Deposits Withdrawals Checks Cashed	Bank Checks Bonds Traveler Checks	Loans: Applications Closings	Night Deposits	New Accounts
B1	229,590	19,126	551	4,348	12,326	15,019	2,161	0	8	0
B3	136,350	6,222	289	2,887	3,777	10,018	262	141	14	0
B4	94,960	6,777	1,781	2,650	4,523	6,755	792	1,305	0	0
B5	153,820	26,600	90	8,022	18,090	0	4,538	0	22	0
B6	101,620	6,038	328	1,644	3,401	7,619	0	296	3	264
B8	125,160	14,164	14	4,488	3,258	5,623	124	0	14	0
B9	61,090	9,591	55	2,809	5,995	2,935	0	0	9	0
B11	177,710	19,021	336	5,747	11,893	11,493	708	0	22	0
B12	100,710	4,122	68	2,408	2,616	7,260	1,869	0	12	778
B13	191,850	31,579	1,307	10,626	19,651	0	2,047	505	24	0
B14	125,460	6,984	383	3,220	4,332	5,653	2,795	93	3	3,500
B15	54,080	4,222	39	1,171	2,896	2,827	602	0	0	0
B18	153,940	11,973	206	6,278	7,425	7,870	974	0	33	0
B20	123,320	16,071	569	4,252	10,301	5,671	1,549	0	11	0
B22	209,700	13,276	422	6,094	3,835	12,017	2,130	64	26	0
B26	83,770	13,206	255	5,309	3,062	3,702	0	0	26	80
B27	184,520	16,304	658	5,107	9,474	12,039	0	0	21	106
B28	170,740	23,127	997	7,707	13,833	1,903	906	436	21	0
B29	135,440	20,713	369	3,662	12,739	9,503	88	0	5	0
B30	499,750	32,294	320	7,447	17,599	23,999	894	0	1	1,931
B31	195,010	21,808	410	6,215	13,408	7,273	1,441	0	14	0
B32	82,720	9,111	370	5,648	6,133	3,966	1,117	48	35	0
B33	623,550	38,758	57	12,378	19,229	47,152	1,589	0	65	4,407

We also calculated the amount of savings by specific resource type the bank could achieve if the 23 less productive branches increased their productivity to the level achieved by the 10 best practice branches (Table 6-2). Specifically, branch B1 should be able to provide its current level and mix of services with 4.5 fewer customer-service full- time-equivalent (FTE) personnel, 1.8 fewer sales service FTEs, 0.3 fewer managers, \$222,928 less in operating expenses, and

1,304 fewer square feet. Of course, these are the DEA results generated from the linear program. In practice, these amounts would be rounded and described as a target level of savings.

In total, the 23 less productive branches identified should be able to reduce customer service, sales service, and managers by 45.5, 45.6, and 8.1 FTEs, respectively. In addition, the expenses could be reduced by about $7 million, and the space utilized could be reduced by about 29,000 square feet. Table 2 also indicates the actual level of resources used and the total resources used by the less productive branches and by all 33 branches. This suggests that the potential resource reductions for this 33-branch system range from 22 percent for customer service-teller FTEs to 33 percent reduction in operating expenses (excluding rent and personnel).

All of these changes would be possible while maintaining or even increasing the volume of services provided by these branches. Table 3 indicates the number of service transactions provided by each less productive branch. It also indicates the added amount of service that could be provided by these branches, in addition to the resource savings suggested in Table 6-2, compared with best practice branches. For example, branch B1 could handle about 15,000 additional deposits, withdrawals, and checks cashed, 2,000 added bank checks, bonds and traveler's checks and eight additional night deposits while reducing the resources needed to attain the efficiency level of the best practice branches. These potential improvements in the resource levels are based on actual performance of other branches rather than simply on abstract theoretical best practice standards. Moreover, DEA gives the benefit of the doubt in measuring inefficiency. This means that it will tend to underestimate the true inefficiency and possible resource savings. Consequently, for some inefficient branches, the actual improvements to productivity may be below the estimate due to unexpected factors preventing change in branch operations. For other inefficient branches, the actual benefits generated can exceed the DEA measure due to the conservative way the inefficiency is calculated.

6.5. IMPLEMENTING THE DEA FINDINGS TO ENHANCE BRANCH PRODUCTIVITY

The DEA information provided a basis for re-evaluating the branch operating ratios and motivated the subsequent use of other techniques to elucidate the work patterns associated with the well-run best practice branches. An analysis of transaction cost data and other operating ratios for best practice and less productive branches substantiated the excess resource utilization located with DEA. These ratios, however, did not take into consideration the differences in service mix provided by the branches. Branches with high transactions per teller were not necessarily the most efficient, as one might think at first. Their inefficiency might arise from lower productivity of the sales service platform personnel and from the fact that their transactions were less complex than some branches with lower transactions per teller.

We conducted field visits and reviews of branch operations, which showed that differences in operations distinguished best practice and less productive branches. For example, we found that the role of the branch manager differed among offices. Using the DEA results, we could focus on differences between the manager's role in best practice branches and in less productive branches. While the branch manager's job was defined consistently for all branches, managers displayed different styles and approaches to managing and rewarding branch employees that resulted in differences in morale and teamwork within branches. Several managers in the best practice branches relied more on cross-training of personnel to handle shifts in demand and had more team based work styles. Also, we discovered notable differences in the use of part-time employees between best practice and less productive branches. Some of the best practice branches had very aggressively recruited part-time personnel to handle peak operating periods. Some of the less productive branches used only full time personnel resulting in more idle time during low service volume periods. We found that more efficient offices had physical layouts that allowed for cross-training of staff and greater ability to balance the workflow.

Some customer services taxed branches unevenly: some were so costly that the bank chose to find alternative ways to provide them. Notably, support costs, including telecommunication services and computer time, varied. As Table 6-2 indicates, branch B33 was reported to have more than $2.5 million of excess support expenses. Initially we believed this to be a computation or data error. Further study led to the surprising conclusion that this excess resource consumption was real and had gone unnoticed because of the size and high profitability of the branch.

Activity analysis provided details of time spent by each employee in each branch (obtained via questionnaires) on such activities as loan solicitation, supervision, training, transaction processing, business development, and inquiries. We tabulated questionnaire to profile the time spent by each type of employee on each activity within each branch. We reviewed and evaluated these distributions relative to the DEA analysis.

This analysis revealed unanticipated differences among best practice and less productive branches. We found that personnel in less productive branches spent more time on tasks for which they were not responsible, or which other professionals could have completed more effectively. For example, the legal issues addressed by branch lending personnel could have been handled more effectively by in-house legal staff. Administrative and technical problem-solving duties consumed considerable time in some branches. This time could have been used more effectively for service transactions and marketing activities.

As a final step, to meet a specific management request, we established optimal staffing models for each category of service branch based on the best practice branches identified by DEA. The bank used these models as benchmarks for existing branches and as models for new branches.

Ultimately, management concluded and reported to the board of directors that the total potential savings identified with DEA were about \$9 million. Actual changes in branch operations that could be implemented within the next year would result in annual savings of over \$6 million of the \$30 million (approximate) operating costs for these branches. (The bank could not achieve some of the savings identified with DEA in the short run, such as square feet reductions). It implemented the changes proposed based on the DEA study. Management realigned each branch's staff according to the model for that type of branch. Simultaneously, it realigned job responsibilities to adopt the approaches used in the best practice branches. This process resulted in substantial staff reductions in most of the less productive offices and also motivated management to review and redefine the objectives of the well-run branches. The bank reduced its total branch staff by about 20 percent within one year after the completion of the DEA analysis. This did require some employee layoffs, a painful and demoralizing process that could have been averted had the bank staffed the branches at the right size initially. The bank hired some of the excess personnel to staff new branches it was establishing.

The branch manager levels suggested by DEA indicated that some branches should operate with less than one manager. All branches had one manager, a level that was considered to be the minimum prior to this analysis. The DEA results motivated reconsideration of this practice. While many of the branches ultimately did require one manager, management had two reactions to this information. In a few branches that were located near each other, it decided have a shared branch manager, creating a satellite branch organization. In cases where this was infeasible for business or geographic reasons, management augmented its expectations of the manager to include added marketing efforts to expand services volume.

The bank moved the managers of strong branches to weaker branches to import their team building and other best practice techniques. The bank redefined the physical layout for new branches: its new leases incorporated office space guidelines that increased flexibility of usage. It centralized some services, which reduced costs and minimized branch responsibility for such services. The bank accomplished these operating changes with *no* noticeable decline in the quality or volume of service.

One interesting result is that one branch that was initially considered to be an average branch based on traditional profit and ratio measures was identified as a particularly strong branch in terms of productivity using DEA. Before we applied DEA, an internal consulting group had visited about half of the branches to evaluate their operations. The group selected what it viewed as the best and the weakest branches and spent several weeks and many travel dollars to complete this review. After we completed the DEA results, the group noted that it had not visited this one outstanding branch. The members observed that had they had these results before their field visits, they could have improved their selection of sites to visit. Management has acknowledged that branch as a model branch. Every individual promoted to branch manager is now required to spend several weeks training at that branch.

6.6. CONCLUSIONS

Growth Bank's objective was to expand it's banking system beyond the initial 33 branches, and management wanted to finance this expansion, in part, by streamlining existing operations. The managers could have used the traditional methods of observation, flow chart process analysis, and ratio analysis, but they were looking for a technique that could provide substantial savings without sacrificing quality of service. Implicitly, they believe all their branches provided service at a quality level that was consistent with their objectives. No explicit analysis of quality was completed. The following chapter described a case where the DEA user was very concerned about maintaining and improving service quality, which requires using a DEA model (Q-DEA) capable of incorporating quality in the analysis.

Using DEA, management located real opportunities to improve operations not apparent with other analytic techniques. The managers did not view problems with data specification and accuracy as confounding the results. If the bank had had data available for other excluded transaction, we might have obtained results that were different in an unpredictable direction, and they might have led to more accurate and detailed insights about branch performance and other ways to improve the branches. While other applications of this type have led to similar results, each application needs to be evaluated based on the available data, the quality of the data, and the potential impact on the results. On balance, this application, its successful results, and management's qualitative assessment of the results vis-à-vis the data used and the way it was measured all suggest that the results provided reliable insights into Growth Bank's branch operations.

An alternate way to use the DEA model is to focus on the service expansion potential while keeping the resource use at the current level. This is particularly interesting in growing markets when management is interested in knowing the extent to which branches can handle increased service volume before needing additional resources. This was of less interest to Growth Bank, as it wanted to expand by increasing the number of branches, financing part of the cost through savings in the existing branch network. This alternative DEA focus provides information about expanding service volume. It is more difficult to influence service volume than to reduce expenses, which is the reason many of the DEA applications focus on cost reductions as a source of increased profitability. However, the use of excess capacity to expand service volume can generate valuable insights about branches that can be expected to grow without added resources and in evaluating the impact of bank mergers (Sherman and Rupert (2005)). Using data envelopment analysis, one can analyze branch operations involving a number of complex transactions with a variety of factors and use that information to pinpoint specific improvements that will result in considerable cost savings. We used DEA to demonstrate that several branches considered the most profitable could be even more profitable because they were not the most efficient and to highlight certain highly efficient practices that were otherwise hidden. Growth Bank achieved $6,000,000 of annual cost savings directly driven by the DEA analysis described in this chapter.

This chapter is based upon Sherman, H.D. and G. Ladino[1], "Managing Bank Productivity Using Data Envelopment Analysis (DEA)" Interfaces, Vol. 25, No. 2, March-April, 1995, pp. 60 – 73, with permission from INFORMS.

[1] George Ladino was Vice President and Director of Global Funds Administration at Citibank N.A.. He previously was Senior Vice President of Fidelity Brokerage Services (FMR Corp.) where he was in charge of a nationwide network of brokerage offices that used DEA to improve branch productivity. He is a C.P.A. and holds a B.A. from Babson College and a M.S. from University of Rhode Island.

Chapter 7

QUALITY-ADJUSTED DEA (Q-DEA)

Incorporating into DEA Productivity Analysis to Seek Low-Cost High-Quality Services

7.1. INTRODUCTION

This chapter provides a roadmap of alternative methods of incorporating quality in DEA benchmarking. If you are managing one of the many services where quality is a critical dimension, the DEA models in this chapter provide productivity improvement techniques that also preserve or enhance service quality. Several methods of incorporating quality into the DEA analysis are described. However, we believe the Q-DEA methodology described in section 7.2 is the most effective option at this time. To illustrate the way Q-DEA can and has been successfully applied and the types of benefits it generates, an application to a U.S. bank branch network and the results of the analysis is described in section 7.3.

Most DEA benchmarking studies do not consider the quality of the services or products (Callen (1991)). Excluding quality can result in adopting perceived best practices that improve efficiency and reduce cost, but *also* reduce quality. Quality concerns are critical to many services and products (healthcare, investment management, home appliances, etc.) but are not addressed in most DEA benchmarking studies. Implicitly, these studies assume that the quality of the benchmarked units is equal, that quality is independent of efficiency, or that quality is not relevant to the analysis. Some may suggest that one can include the quality measures into DEA. The numerical example in the next section, however, shows that the DEA efficiency usually improves as the number of inputs and outputs increases. In fact, increasing the number of inputs or outputs increases the number of constraints in the envelopment model and increases the number of variables in the multiplier model. In either case the optimal value (= efficient score) cannot decrease and generally will increase. Adding quality as an output results in diminished insights about how and where productivity can be improved. There are situations where this is acceptable and where this is the preferred result. Such situations will be covered in this chapter.

We further illustrate that including quality may suggest a tradeoff between quality and productivity that can be contrary to management objectives and may suggest unacceptable changes in operating methods. For example, spending more time with a customer can result in higher service quality and will reduce the productivity due to the added time servicing the customer and management may encourage this type of tradeoff. However, if the service provider increases his/her productivity by spending less time with the customer resulting in a low customer satisfaction quality rating, this tradeoff may be contrary to management's objectives of retaining customers to ensure continued demand for the business services. Here productivity improvement in the short run can lead to lost customers and reduced long-term profitability

The need to consider quality in benchmarking can also be seen in bank branch examples discussed in previous chapters. Without explicitly including quality in the branch analysis, DEA could identify best- practice branches that use fewer resources even if they achieve this by providing low-quality service. Emulating low-cost/low-quality branches could result in loss of valuable customers and reduced profitability over time, a result that bank managers are unlikely to find acceptable. Hence, if we seek high-quality low-cost best-practice benchmark branches, quality needs to be considered. The best practice ERS units located with DEA without including quality may be efficient and low quality, which is likely to be a sub-optimal definition of best practice, which would not be acceptable as a model in a service where quality is a critical dimension. Incorporating quality into DEA will enable this methodology to identify best practice units that are highly efficient *and* high quality.

This chapter explores the way DEA can incorporate quality in its benchmarking to help identify high-quality and high-efficiency best-practice benchmarks[1]. Several alternative DEA benchmarking methods that consider quality are considered, including ones proposed in other DEA studies. A new method designated Q-DEA is suggested as more comprehensive and the results of applying Q-DEA to a bank branch network are reported. The advantages and weaknesses of Q-DEA are identified, providing a potential basis for developing more comprehensive and effective DEA benchmarking methods.

Section 7.2 describes several methods of considering quality in the DEA analysis; each is illustrated through a benchmarking application to a simplified data set. The advantages, limitation, and assumptions about quality are explained to allow managers to adopt the method most compatible with the way quality is managed in their organizations. Section 7.3 describes the results of applying the suggested Q-DEA evaluation to a multi-state bank branch network, the reactions of management to these results, and the impact of implementing the results on the bank's performance. The benefits and limitations of using Q-DEA for benchmarking are reviewed. The final section 7.4 identifies several key issues where additional research on incorporating quality in DEA warranted.

7.2. INCORPORATING QUALITY INTO DEA BENCHMARKING

A simplified version of the real data set in section 7.3 is used to illustrate the impact of quality on the DEA productivity benchmarking. The example, presented in Table 7-1, is designed to illustrate and contrast DEA benchmarking results using several different analytical models. Specifically analysis of this data set will (i) illustrate the impact of excluding quality; (ii) illustrate the benefits and weaknesses of incorporating quality as suggested in several DEA studies; and (iii) describe the advantages and limitations of a new benchmarking approach that adjusts DEA for quality, referred to as Q-DEA.

The example in Table 7-1 has the following characteristics (a graph plotting resources used by each branch is given in Figure 7-1):

(i) Each branch processes the same number of identical transactions. This simplified assumption is naturally expanded to multiple transactions in the application in Section 7.3. Although DEA

[1] High productivity and low cost are used interchangeably, suggesting that efficient resource use reduces costs. A technically efficient unit may not be low cost if it does not pay competitive prices. The price issue will not be considered in this chapter.

accommodates multiple outputs in natural units, we use one type of output and one output level to facilitate graphic illustration of the approach.

(ii) Branches use two types of resources: teller (labor) hours to complete the transactions and non-personnel expenses such as costs of supplies, utilities, etc., to support the tellers completing the transactions. The amount of hours used and expenses are reported for each branch.

(iii) The quality of service is measured for each branch. One typical service quality measure is a mystery shopper score based on visits by artificial customers employed by survey companies. These artificial customers complete a questionnaire measuring service on predetermined quality measures and provide an overall rating for the services. In banking, they would rate each teller and the entire branch. The highest quality rating is generally set at 100. In this example, Branch I has the lowest quality rating of 60.

(iv) Branches A through E use the identical levels of resources as Branches G through K. For example, Branch A uses the same resources as G (20 hours and $300 of other expenses). Similarly, B and H, C and I, D and J, and E and K have the same resource use and the same transaction levels. The only difference within these branch pairs is their quality level. For example, Branch A has a quality rating of 100 while Branch G has a quality rating of 90. Hence, they are equally productive but Branch A has higher service quality and would be considered a better model branch than Branch G.

(v) All the data are assumed to be accurate and represent the appropriate measures of quality, transactions, and resources used.

Table 7-1. Bank Branch Example

Branch	Transactions	Quality (maximum = 100)	Teller Hours	Non-personnel expenses
A	1000	100	20	300
B	1000	100	30	200
C	1000	100	40	100
D	1000	100	20	200
E	1000	100	10	400
F	1000	80	20	150
G	1000	90	20	300
H	1000	70	30	200
I	1000	60	40	100
J	1000	90	20	200
K	1000	80	10	400

Branches A & G, B & H, C & I, D & J, E & K use the same resources. The quality scores among these pairs of branches differ.

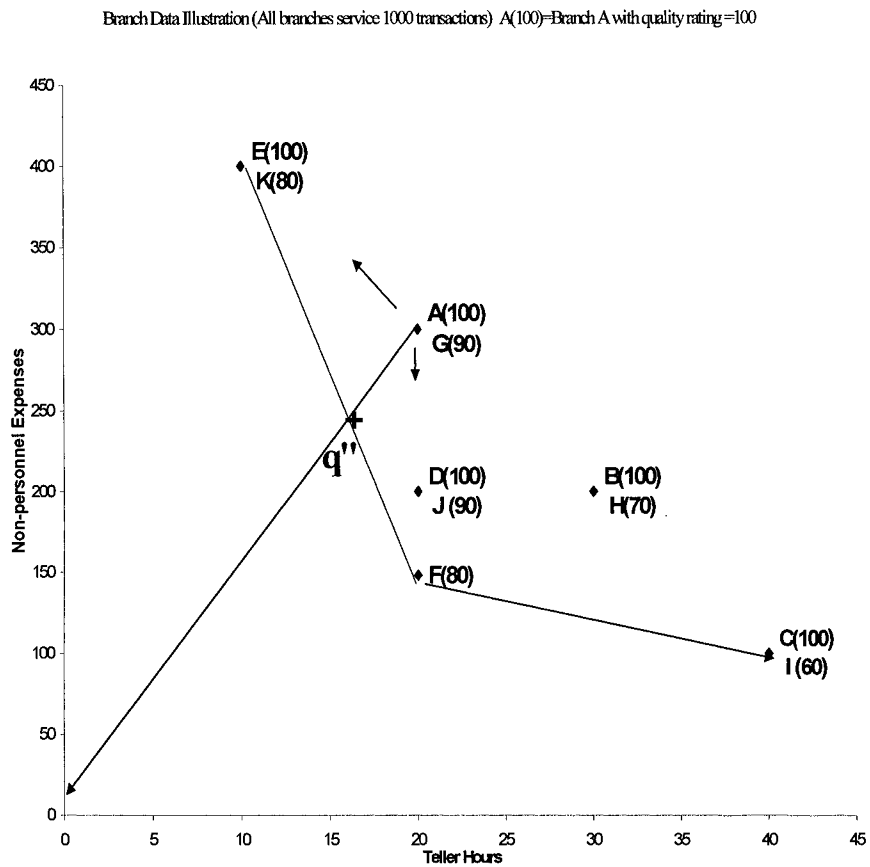


Figure 7-1. Branch Data Illustration

Questions a manager of this bank might ask include: How are these branches performing? Which are the best-practice branches? Could the network performance be improved if all branches emulated the best-practice branches? DEA-based benchmarking is one way to begin to address these questions.

Four DEA models are considered in their ability to consider quality in the productivity analysis: Model I, basis DEA, Model II – including quality as an output, Model III – treating quality as an independent dimension, and Model IV – quality adjusted DEA (Q-DEA).

7.2.1 Standard DEA Model (Model I)

If one applied the standard DEA model, e.g., the CRS model (2.1), to benchmark these branches using the approach found in many DEA studies, one would run a model with one output (transactions), and two inputs (hours and other expenses). The result would be a DEA rating where the highest productivity rating is 1, signifying best-practice. A rating below 1 signifies that the unit is inefficient compared to other units in the study.

Table 7-2. Model I - Benchmarking Productivity with DEA *Excluding Quality*

		Potential resource reductions		DEA benchmark reference set			
Branch Name (quality rating)	DEA benchmark productivity rating	Teller Hours	Non-personnel Expenses	Reference Branch (quality rating)	Reference branch weight	Reference Branch (quality rating)	Reference branch weight
A (100)	0.81	4	56	E(100)	0.375	F(80)	0.625
B (100)	0.73	8	55	C(100)	0.091	F(80)	0.909
C(100)	1	0	0				
D(100)	0.93	1	14	E(100)	0.143	F(80)	0.857
E(100)	1	0	0				
F(80)	1	0	0				
G(90)	0.81	4	56	E(100)	0.375	F(80)	0.625
H(70)	0.73	8	55	C(100)	0.091	F(80)	0.909
I (60)	1	0	0				
J(90)	0.93	1	14	E(100)	0.143	F(80)	0.857
K (80)	1	0	0				
Total potential resource reduction		27	250				

The results of this DEA benchmarking analysis are presented in Table 7-2, where Branches C, E, F, I and K are the best-practice benchmarks, indicated by the DEA benchmark productivity rating of 1. Note that branch F is a benchmark for all inefficient branches. The fact that F has a lower quality rating (80) than A and B (each is 100) is not considered in this analysis. If management were to have branches emulate the efficient Branch F, the network might adopt methods that reduce operating costs but also lower the branch service quality, which could ultimately be costly and dysfunctional.

A key characteristic of DEA as a benchmarking tool is the following: If any inefficient branch (DEA rating < 1) adopts best-practice branch methods *without* damaging quality, that branch would reduce its resource use and operating costs, improving its profitability regardless of the cost per teller hour.

Model I in Table 7-2, frequently found in DEA studies, illustrates one limitation in this DEA analysis: quality is not considered. Any effort to use this information without considering quality has the potential to reduce service quality, which may have significant effect on future costs.

Few DEA studies explicitly address quality and those that consider it have not fully adjusted for quality. For example, one DEA study of a bank-branch network in Greece, Athanassopoulos (1997) considers the relationship of the DEA productivity scores with quality. An independent quality measure is developed based on customer surveys and the statistical relationship between quality and the outputs in the DEA model, such as new accounts and loans, is measured. This suggests the extent to which quality is impacting sales of services. The DEA scores are calculated as above - model I - with *no* quality adjustment. A concept of effort effectiveness is proposed that would be a function of efficiency as measured with DEA and quality. This study does not actually combine these two key elements into this effort effectiveness measure. Rather, it acknowledges the void and calls for research to find ways to combine quality and efficiency – an objective of this chapter.

7.2.2 Quality as an Output in Standard DEA Model (Model II)

We now demonstrate that directly adding quality as a DEA (output) measure does not help in discriminating the performance. If we include the quality measure as a second output in the DEA Model I, the efficiency of each branch will be improved, as evidenced in Table 7-3. Note that Branch D becomes efficient. It is theoretically true that the DEA efficiency score will not be decreased if additional output(s) and (or) input(s) are included.

Branches F and I are designated best-practice branches with model I and II even though they have low quality scores of 80 and 60, respectively. For some applications, this result may be problematic. The lower service-quality rating of 80 was not low enough to disqualify Branch F as a best-practice branch. Here, high productivity compensates for low quality. Should the manager of the service units being benchmarked accept this or any tradeoff between quality and productivity? If this type of quality-productivity tradeoff information can be identified by management or from other reliable sources, then we can use the cone ratio DEA model (Charnes et al. 1990) or other weight restriction models discussed in chapter 5 to refine the DEA efficiency[2]

Table 7-3. Benchmarking with Quality as an Output

		DEA benchmark reference set					
Branch Name (quality rating)	DEA benchmark productivity rating	Reference Branch (quality rating)	Reference branch weight	Reference Branch (quality rating)	Reference branch weight	Reference Branch (quality rating	Reference branch weight)
A (100)	0.86	D(100)	0.714	E(100)	0.286		
B (100)	0.86	C(100)	0.286	D(100)	0.714		
C(100)	1						
D(100)	1						
E(100)	1						
F(80)	1						
G(90)	0.82	D(100)	0.143	E(100)	0.357	F(80)	0.5
H(70)	0.73	F(80)	0.901	I(60)	0.091		
I (60)	1						
J(90)	0.96	D(100)	0.417	E(100)	0.083	F(80)	0.5
K (80)	1						

Service quality is a complex dimension in itself, however, and its importance varies in different industries. What should the tradeoff be between productivity and quality when selecting the surgeon to do a cardiac-bypass operation? Does the tradeoff between cost and quality of the physician change for less-severe types of medical care? What is the cost/quality tradeoff in selecting a public or private school? It is very difficult to quantify this type of tradeoff information.

On the other hand, quality/productivity and quality/cost tradeoffs are not readily acceptable in many applications. Even in this example, would one want to ask Branches A and B to emulate lower-quality Branch F, which could mean reduced costs but also reduced service

[2] One study that includes quality as an output in the context of airplane maintenance is Rouse, Putterill and Ryan (2002).

quality? Would a manager advocate reducing service quality? Is there a clear enough measure of the value of service quality to be able to quantitatively define economically attractive tradeoffs?

7.2.3 Independent Quality and Productivity Dimensions (Model III)

One model that avoids an automatic quality/cost tradeoff is to treat quality as a dimension independent from productivity and benchmark on these two dimensions simultaneously. The best-practice would be defined as high quality low cost or high quality high productivity (HQ-HP). This approach, we refer to as Model III, maps the benchmarked units in terms of quality and productivity. The portrayal of these branches in Figure 7-2 reflects the way quality and productivity are addressed in the other three DEA studies that consider quality, by Bessent et. al. (1984) in public schools, by Chilingerian and Sherman (1990) in physicians, and by Soteriou and Zenios (1999) in banks.

This chart can be separated into segments to reflect the way a manager chooses to define high quality and high productivity. For example, in Figure 7-2, the graph is broken into quadrants: high quality and high productivity (HQ-HP), low quality and low productivity (LQ-LP), low quality and high productivity, and high quality and low productivity. The advantage of model III is that it identifies branches in the HQ-HP quadrant that can serve as best-practice benchmarks for other branches. It culls out Branch I as the branch that is efficient but provides the lowest-quality service. The quality and productivity dimensions are measured independently in this analysis.

To identify the most appropriate benchmark branches, management would establish a range of quality and productivity that meets its operating objectives. In this example, quality levels of 90 and above are marked as acceptable benchmarks and productivity levels of 0.90 and above are marked as acceptable. This would identify Branches C, E, D, and J as best-practice benchmarks that are considered high-quality/high-productivity branches. Branches F, K, and I are no longer benchmark branches, in contrast to Model I where these were best-practice branches. These three branches might be challenged to improve their service quality even if this required added resources. Similarly, Branches G, B, and A might be expected to improve productivity without sacrificing their high service quality. Model III incorporates quality and does not allow for the quality/cost trade-off seen in Model II where quality is included as an output. For example, Branches F and K had high ratings with Model II because their productivity was high enough to offset and possibly disguise low quality that is apparent in this analysis. Model III singles out Branch H as the weakest performer, the only branch in the low-quality low-productivity quadrant and one that would likely be expected to improve both quality and productivity to move to the HQ-HP quadrant.

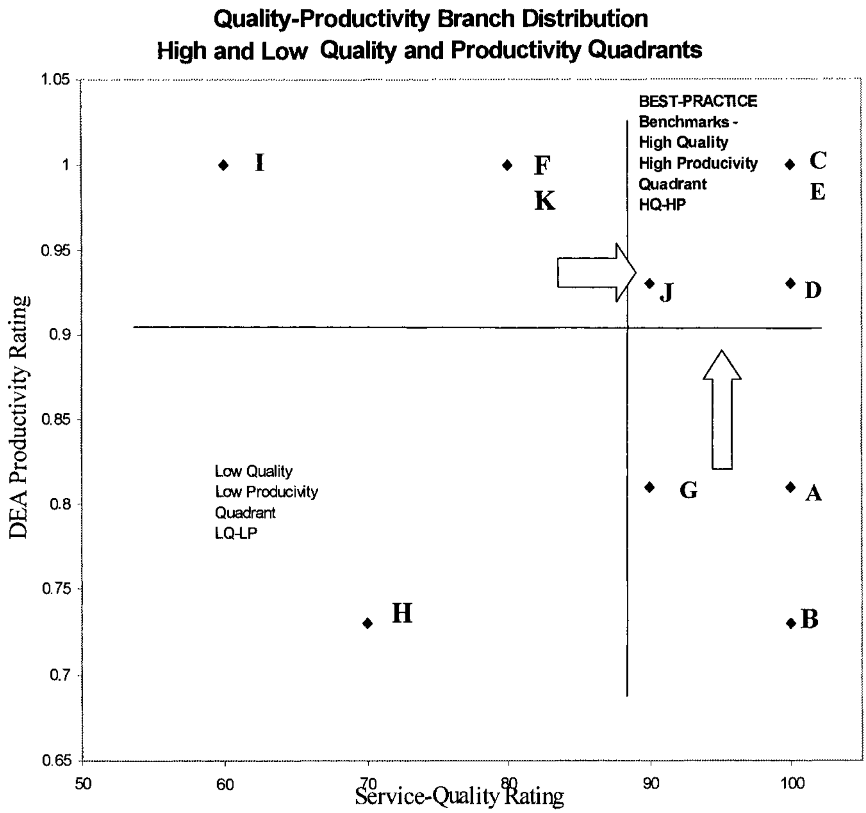


Figure 7-2. Model III

Model III has the advantage of making quality/productivity tradeoffs more visible and manageable. It allows management to identify benchmarks that will move the organization to higher quality and productivity.

Note that Model III includes low-quality/high-productivity branches in the set of reference branches which *under-values* the productivity of the other branches and overestimates the amount of resource savings that may be possible as the branches move to the best-practice benchmark levels. Adjusting the DEA productivity analysis to remove the impact of low-quality branches is the objective of Model IV - designated quality-adjusted DEA (Q-DEA).

7.2.4 Quality-Adjusted DEA (Model III)

To filter out the impact of low-quality branches, a multi-stage DEA analysis is proposed, essentially reapplying DEA to an adjusted data set to achieve a particular objective. The objective here is to develop a DEA analysis where all the most productive branches (DEA rating of 1) are also high-quality branches, resulting in a set of high-quality/high-productivity benchmarks and a target productivity level that is not skewed by low-quality/high-productivity units.

We develop the following algorithm to execute the Q-DEA:

Step 1: Run DEA Model I.

Step 2: If the number of high-productivity/low-quality (HP-LQ) units = 0 then stop. Otherwise, remove the HP-LQ units and go to step 1.

Table 7-4. DEA Productivity Ratings

Branch Name	Quality	DEA unadjusted for quality (Model I)	Quality Adjusted DEA (Model IV)	Branches with increased productivity rating with Q-DEA
A	100	0.81	0.86	←==
B	100	0.73	0.86	←==
C	100	1	1	
D	100	0.93	1	←==
E	100	1	1	
F	80	1	Low Quality	
G	90	0.81	0.86	←==
H	70	0.73	0.86	←==
I	60	1	Low Quality	
J	90	0.93	1	←==
K	80	1	Low Quality	

High-productivity branches with low service quality identified in Model I are removed from the data set and DEA is rerun to determine if the new high-productivity benchmark branches are also high quality. If any of the high-productivity branches in Model IV, Branches D, C, E, and J, were low quality, they would be removed and the Q-DEA benchmarking would be rerun until this process generated a set of branches with a DEA rating of 1 and high-quality scores.

In the branch illustration, the quality-adjusted DEA rating is achieved by removing the branches with high DEA rating of 1 and unacceptable quality levels. In this case, Branches I, F, and K are eliminated from the analysis, as they have DEA ratings of 1 but are below the minimum quality standard, in this case below 90. The result of this analysis is reflected in Tables 7-4 and 7-5.

Table 7-5. Q-DEA Benchmarking

		Potential resource reductions		DEA benchmark reference set			
Branch Name (quality rating)	DEA Productivity Rating	Teller Hours	Non-personnel Expenses	Reference Branch (quality rating)	Reference branch weight	Reference Branch (quality rating)	Reference branch weight
A (100)	0.86	3	43	D(100)	0.714	E(100)	0.286
B (100)	0.86	4	29	C(100)	0.286	D(100)	0.714
C(100)	1	0	0				
D(100)	1	0	0				
E(100)	1	0	0				
F(80)	Low quality						
G(90)	0.86	3	43	D(100)	0.714	E(100)	0.286
H(70)	0.86	4	29	C(100)	0.286	D(100)	0.714
I (60)	Low quality						
J(90)	1	10	0				
K (80)	Low quality						
Total potential resource reduction		24	143				

Table 7-4 indicates that by removing the low-quality/high-productivity branches from the analysis, the remaining inefficient branches have higher productivity scores, and Branches D and J become best-practice branches. Each branch on the new best-practice frontier line now provides service that is above the target quality level of 90, i.e., all benchmark best-practice branches are high quality and low cost. Less-productive branches are identified based on a comparison with high-quality/low-cost benchmarks. Hence, a manager of a branch with a low score cannot question the rating on the grounds that the reference set of benchmark branches are sacrificing quality to achieve lower costs. Bank management can reasonably question why branches with low Q-DEA benchmark scores cannot achieve the best-practice levels.

Table 7-5 indicates the expected savings and the benchmark reference branches for each branch in the Q-DEA analysis. Branch A is now identified as having a productivity rating of 0.86 compared with benchmark branches D and E, both of which have quality scores of 100. This is in contrast to Model I, where branch A was compared with Branches F and E, and where Branch F had a lower quality rating of 80 (see Table 7-2). In addition, the expected resource savings are measured based on the HQ-HP benchmarks. In this case, Q-DEA suggests that Branch A has the potential to reduce teller hours by 3 hours (versus 4 with no quality adjustment) and reduce expenses by $43 (versus $56 with no quality adjustment). The objective is to achieve lower-cost operations by reducing excess resources while maintaining or improving quality through benchmarking. The total potential savings suggested by the Q-DEA analysis for this branch set is 24 teller hours and $143 non-personnel expenses (Table 7-5), which is lower than the amount suggested with no quality adjustment in Model III (27 teller hours and $250 non-personnel expenses in Table 7-2). In this example, with only 11 service units or decision-making units (DMUs), the Q-DEA algorithm offers new insights into which units have high productivity and high quality. However, this data set is designed to make these differences clear. Applying this algorithm to small data sets may result in less significant or even no added insights.

7.3. Q-DEA BENCHMARKING APPLICATION TO A BANK BRANCH NETWORK

A U.S. bank with over 200 branches distributed over five states was interested in rationalizing their branch network and reducing operating costs by adopting best-practices among their branch network. Prior to this study, the branch performance was evaluated within regions. No branch wide benchmarking was in use and management of branches was primarily the responsibility of regional managers. This study transpired at a time when there was pressure on the bank to reduce operating costs to boost earnings both to satisfy investment analyst demands and to maintain a strong position for acquisition of other banks. Service quality was considered very important due to intense competition in most of their geographic markets and the value placed on the branch network in retaining customers, generating fees, and supplying low-cost demand deposits. Having read publications about this technique, the network manager wanted to evaluate the possible use of DEA as a benchmarking tool to reduce operating costs.

Banks in the U.S. generally have three types of personnel in a branch: tellers, platform, and management. Tellers handle most high-volume standard transactions including deposits, withdrawals, and bank checks. Platform personnel handle more complicated and customized transactions such as loan applications, new and closed accounts, wire transfers, and retirement (IRA) and other investment accounts.

This bank had been using a staffing model that assigned tellers based on total teller transactions and peak demand periods. It was not sensitive to the mix of transactions. In addition, the staffing model applied only to teller transactions and did not consider total resources used by the branch. Another serious limitation of the staffing model was that it ignored the activities of branch management and platform personnel in the branch. The objective was to use DEA benchmarking to uncover new ways to improve branch operations.

The information systems of the bank were relatively advanced and captured branch transactions in detail. Branch performance measures included a profit-and-loss statement that focused primarily on expenses, net interest earned on deposit, and fees generated by the branch. Quality of service was measured based on evaluations by mystery shoppers posing as customers. The maximum score was 100. The issues of customer retention and limitations of mystery shopper scores were well understood by branch management. They believed that these mystery shopper scores were a good measure of quality and approved of this method of service evaluation. . Other measures of service quality in banking are discussed in Athanassopoulos and Giokas (2000).

To develop the DEA model that would be used to benchmark the branches, the regional branch managers were asked to define the transactions handled by their branches, to indicate branches that had unusual activities or characteristics, to outline the types of resources used by the branches, and to identify branches that had unusual working hours and high-demand periods. Based on this information, the inputs and outputs to be used were established as summarized in Table 7-6. Various branches were excluded from the analysis because they were new, closing, or were special branches designed to handle unusual transactions or transaction volume. The first DEA analysis began with data on the customer-service quality, four inputs including personnel and operating expenses, and eight types of transaction outputs covering teller and platform activities for 225 branches.

Table 7-6. Branch Data Used for Q-DEA Benchmarking

Inputs:	**Outputs:**
1. Platform Full-Time Equivalents (FTEs) 2. Teller FTEs 3. Management FTEs 4. Postage, supplies, telephone, travel expenses	1. Deposits, withdrawals, checks cashed 2. Bank checks 3. Bond transactions 4. Night deposits 5. Safe deposit visits 6. New accounts – time and demand deposits, Certificates of deposit, IRA, Safe deposit boxes 7. Mortgage and consumer loans – applications, closings 8. Automatic Teller Machines – serviced by branch
Quality Measure	
Mystery Shopper Scores – branches below 90 out of maximum score of 100 were deemed disqualified to be best-practice models (cutoff was approximately the network average score)	
Branch classification schemes: Urban, Rural, Suburban Commercial, Personal Transaction volume – size classification Deposit dollar – size classification Vacation, retirement, community development (low-income urban), supermarket	**Transactions weights** were used to increase the power of the DEA analysis.

Quality of service was a key issue, as management intended to challenge low-scoring branch managers to move to best-practice operating levels found in other branches or to justify their existing resource levels. The best-practice branches that would be the basis for this challenge needed to be high quality and high productivity (low cost). The idea was to move all branches to the high-quality/low-cost best-practices in the branch system. Two characteristics made the use of Q-DEA appear particularly appropriate beyond its ability to incorporate quality: 1) The geographic separation and regional organization meant that some of the benchmark branches would not be familiar to other branch managers; and 2) There were too many branches to allow management to qualitatively assess the interactions of quality and cost of operations. The DEA model used was the CRS DEA model, which does not adjust for economies of scale. This means that large branches can be best-practice benchmarks for small branches and vice versa. There is no strong external evidence that branches have increasing or decreasing returns. To determine whether this was a potential issue in practice, separate DEA analyses of different branch types and sizes were completed before the results were used to influence branch operations. The only place where this altered the results was for small branches, where there are minimum staffing levels that are required to meet basic operational controls. The productivity ratings and target resource levels for small branches used for the small-branch analysis was used in guiding branch managers on the improvement targets. The small-branch analysis also leads to devising a new model for the efficient high-quality branch (O'Keefe (1994)).

Phase 1: Improve branch network quality

The data on 225 branches were accumulated and run with the inputs and outputs in Table 7-6 using the multi-stage Q-DEA process described in Section 2 of this chapter. The input measures were the average number of FTEs for each type of employee and total expenses for the quarter and the output measures were the number of transactions for the quarter. Management was asked to select a cutoff point to disqualify benchmark branches with unacceptable quality. The average quality of the branch system was about 88 , and management determined that any branch with quality scores below 90 would be disqualified as a benchmark branch. The Q-DEA result after several iterations culling out low-quality high-productivity reference branches was that 32 branches were identified as highly efficient but with low quality. If these were allowed to serve as benchmarks, there was the potential that the low-scoring branches would be challenged to achieve a target operating resource level based on branches that provided low quality service.

The initial reaction of the network manager was that this was a surprisingly large number of disqualified low-quality branches. His view was that these branches needed to improve their quality and be re-incorporated into the analysis. Essentially, the use of the benchmarking to reduce costs was deferred to deal with service quality, which was a higher priority. A program to focus on improved service quality was implemented and the branches were asked to focus on improving service quality.

Phase 2: Use Q-DEA to reduce branch network operating costs

Benchmarking was resumed about half year after the program on quality improvement was initiated. Another data set was developed for a subsequent quarter using the same data elements with 229 branches. The net increase in branches reflects some branches closing and others that were just opening during Phase 1 but were fully operating and included in the second analysis. Average quality of the network increased to about 90 and branches with DEA efficiency scores of 1 with quality below 90 were disqualified as best practice branches. The second Q-DEA analysis identified only eight branches that had low quality/high productivity. Management considered this to be more acceptable as a starting point for the benchmarking analysis and separately asked that these eight branches be studied to determine if their operating cost levels were too low and were causing the low service quality.

At this point, the bank had a listing of the best-practice branches in terms of high-quality and low-cost operations. The breakdown of branches is summarized in Table 7-7.

Table 7-7. Q-DEA Benchmarking Applied to a US Branch Network

	Phase 1: Raise Quality			Phase 2: Increase Profitability		
	# of branches	Average quality rating	Profitability ranking	# of branches	Average quality rating	Profitability ranking
Best practice	60	90.5	2	46	92.1	2
Less productive	133	89.1	3	175	90	3
Low-Quality/ High-Productivity	32	80.7	1	8	77.4	1
Total	225			229		

Note that the best-practice branches are more efficient, more profitable (partly due to efficiency) and have higher-average quality than the less-productive branches. The low-quality/high-productivity branches have the lowest average quality but are also the most profitable, possibly because personnel resources are not sufficient to provide high-quality service. Management's reaction to Phase 1 was to focus on raising service quality. Phase 2 begins about half year after phase 1 and quality scores are higher with fewer low-quality/high-productivity branches.

For each branch, the following information was generated: the DEA rating, the target resource savings, the benchmark branches that comprise the DEA efficient reference set of branches, and the weight assigned to each branch. This is similar to the information in Table 7-5. The Q-DEA results identified 46 best-practice high-quality/high-productivity branches, 8 low-quality/high-productivity branches and 175 branches with DEA ratings less than 1.0 that had potential to reduce costs by adopting best-practice operating methods. Of the 175 branches with some inefficiencies, 147 had DEA ratings at or below 0.9 (see Table 7-8).

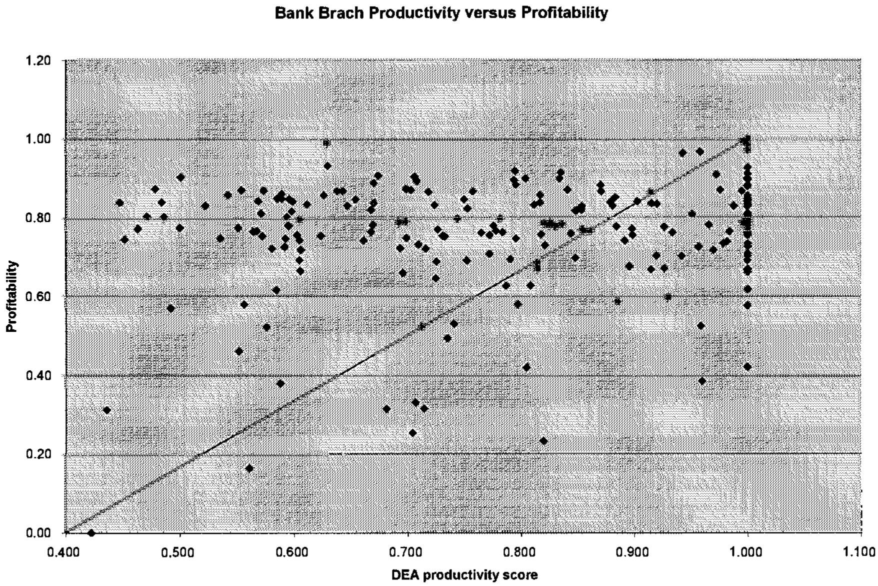
Table 7-8. Q-DEA Benchmarking Distribution of Productivity Ratings in Phase 2 in the U.S. Bank Application

Q-DEA Productivity Rating	# of Branches	
1	46	Best Practice – High-quality/High-Productivity
		Less-productive Branches
.90 -.99	28	
.80 - .89	41	147 Branches
.70 - .79	37	with ratings below .90
.60 - .69	27	
.50 - .59	30	42 branches with ratings below .6
.40 - .49	12	Suggests that they are using 40% more resources than best-practice benchmark branches.

Various tests and analyses were completed by branch management to ensure that these results appeared to make sense. The 45 best-practice branches were reviewed to determine if there were any characteristics that would make them abnormally efficient or poor benchmarks. This included reviewing the results with regional managers and soliciting their qualitative reaction as to whether these were considered model branches. The least-efficient branches were also reviewed to determine if these were already known to be poor performers and whether there were characteristics that would handicap these branches.

The initial reaction was that Q-DEA provided a perspective that differed from the methods already in use and that the results were surprising in many ways. Examples of surprises in this initial stage included the fact that the 46 best-practice branches included a wide variety of branch types. Small branches, vacation community, low-income urban, and retirement community branches were among the best-practice branches. These branch types were thought to be potentially the least efficient because of seasonal staffing requirement, slower transaction times, low-income and multi-lingual customers, and difficulty in staffing small branches due to the need for a minimum staff level to maintain adequate financial controls over branch transactions. Similarly, many of the largest branches were identified as being inefficient and several of the branches that were considered to be the best performers in terms of profitability were among the lowest-rated branches.

The following chart shows a plot of profitability vs. efficiency score.



There is no clear correlation between profitability as measured by the bank and productivity, which illustrates the ability of DEA to locate excess operating costs that are not visible from the accounting measures or statistical regressions. The primary reason is that the bank measures branch profits based on revenues from fees and attributed net interest income from deposits generated by the branch less operating expenses. The operating costs are also included in the DEA analysis but these costs are compared with the activities of the branch in servicing customers such as deposits, visits to safe deposit boxes etc. and these are not reflected in the profitability. This is also the reason DEA complements the accounting profit analysis - both used together are most effective.

While these results were surprising, scrutiny of the data supported the Q-DEA results and management decided to proceed to the next step.

The Q-DEA results were distributed and explained to each regional manager in charge of a group of branches. The meaning of each element and the interpretation on a branch-by-branch basis was explained. The regional managers then met with each branch manager that had a DEA score of 0.90 or lower to focus on branches where the greatest excess resources are likely to be located. The branch managers were asked to consider the analysis, compare their operations with branches identified as best-practice benchmarks for their branch, and determine if and how they could alter their operations to move toward the types of branch procedures, staffing, and operating costs that characterize the best-practice benchmarks. One can visualize this process by revisiting the example in Figure 7-1. Essentially, the managers of Branches A and G (see Figure 7-1) were challenged to reduce their resources while maintaining or improving their quality by moving toward Branch E, D, or J, or toward the point "q" based on what appears feasible for each branch manager knowing their business environment. This information and challenge was provided to each of the 147 branches with Q-DEA scores below 0.90.

Results of Q-DEA Benchmarking

Management used the Q-DEA results to focus branches on identifying ways they could reduce operating costs to the best-practice levels. To rely on these analyses, management needed the confidence that there were real best-practice benchmark branches in their network that provide the same or higher-quality service with fewer resources than branches with low DEA scores. They used this to challenge the other branches to improve their performance. *If the results were not quality adjusted, management of this bank would not have used these results.*

The Q-DEA process generated substantial cost savings, a wide variety of responses and several centralized initiatives. Some of the responses are in reaction to very specific findings reflected in the DEA results that the bank had not identified with other performance measurement methods in use. Examples of the findings are described below.

The potential and actual resource savings from the Q-DEA benchmark analysis are summarized in the Table 7-9.

FTEs were reduced by 149, or over 7% of the network staff. These reductions were all generated from the branch level by the branch managers working with their district managers (in charge of a group of branches.) The changes in staffing were based on their analysis of the Q-DEA benchmark results. Savings were generated among all branch types including high- and low-deposit branches, and high- and low-transaction volume branches. The reductions realized were all believed to be feasible by the branch managers but were not identified before the Q-DEA benchmarking. Branch managers were using benchmarks to reduce excess resources. These savings are believed to be more likely to endure and less likely to damage service quality compared with more common across-the-board staff reductions. The endurance and long-term impact on service quality were not measured.

Table 7-9. Potential and Actual Savings from Q-DEA

	FTEs	% of all branch FTEs	Non-personnel expenses	% of all branch Non-personnel expenses
Potential Resource Savings (Q-DEA benchmark Rating < 1)	422	21%	$1.3 million	28%
Actual Savings (within 6 months of Q-DEA analysis)	149	7.4%	Not measured*	

*The largest expense impact was the identification of high telephone charges in two of the state branch networks compared to telephone charges in other states. This comparison was never made until the Q-DEA benchmarking. In one of these high cost states with a system of over 30 branches, new phone arrangements reducing their phone system costs were negotiated triggered by this analysis. In the other high-cost telephone state, there were fewer than 20 branches and no changes in the phone system resulted.

One group of branches where management was aware that Q-DEA potential staff reductions exceeded 60 FTEs reported that they could reduce their staff by only 6 FTEs. The regional manager reported that added savings were not possible due to changing business conditions and the need for these resources to build their business. *This was not challenged* even though it was not likely to be accurate, in part because these branches had the reputation as being the most profitable. Q-DEA analysis clearly identified significant excess resources in these branches. Bank politics prevented further challenge of this decision, though there was recognition that if the business growth did not materialize, the staffing would be re-evaluated.

Small branches: Of particular interest was the set of small-branch best-practice benchmarks. The bank was considering closing smaller branches and based on these findings began to develop a new small-transaction branch model that incorporated the part-time and flexible hours that the best-practice small branches had experimented with. A separate Q-DEA analysis was run benchmarking only small branches to refine the small-branch model being developed based on small best-practice branches. One of the senior bank managers was interviewed in a bank trade magazine and he commented on this Q-DEA process *before* the above savings were achieved. He noted that "... we found there's some very small productive branches, and we found a lot of large branches that were not very profitable."... "Many banks have small branches and they are looking for solutions other than just closing them" (O'Keefe (1994)). The way small branches were measured was altered as well. In addition to the deposit size, transaction size was considered. Much of the excess FTEs from small branches were found in small branches with low transaction levels that previously were automatically given the same FTEs as other small branches.

Large-deposit and high-transaction branches: Another bank wide response to Q-DEA benchmarking was the recognition that large branches with high deposits appeared very profitable but were using more than 1/3 of the excess FTEs located in the study, suggesting that they should be even more profitable.

Part-time FTEs: Several best-practice branches were able to handle transactions efficiently with high quality service by allowing tellers and platform personnel to assist others during peak periods or when individuals were absent *and* by aggressive use of part-time personnel. It was determined that many low-productivity branches were not using part-time personnel and that this was due to limited ability to attract and hire this type of employee. The bank subsequently adopted a program to provide enriched health plans and other benefits to part-time employees to enable these branches to hire and retain part-time personnel with the objective of operating with a more flexible staff, resulting in a net decrease in FTEs.

Branch-by-branch changes: There were numerous other changes that occurred from individual branch analyses. Two examples of the kinds of insights generated from these benchmarks that had formerly escaped the attention of the branch managers, regional managers, and network management follow.

Two branches in geographically distant regions were compared based on the

Q-DEA reference branch listing. It was noted that one branch had excess FTEs but was otherwise similar to the best-practice branch. The branch managers discussed their operations and determined that they both had drive-up teller windows that were detached from the branch. The best-practice branch had converted this to a drive-up ATM to eliminate the need to staff the drive-up window and this was received well by their customers. The low-rated branch had not considered the idea. Following this analysis, the low-rated branch changed its drive-up window to a drive-up ATM.

One low rated branch reflected no excess FTEs but excess operating expenses. The regional manager could not explain this at the introductory meeting. After meeting with the branch managers and investigating this expense, the bank identified billing errors dating back over one year. The branch was paying for its phones and another phone line in the same building that was used by another business. Among the 147 inefficient branches, there were many similar findings unique to individual branches.

7.4. CONCLUSIONS AND FUTURE RESEARCH

Many of the DEA applications are in services such as education and healthcare where quality is at least as important, and probably much more important than in the bank application. Even in the bank application, if service quality had not been explicitly included in the analysis, management would not have used the findings. The value of using DEA as a tool to improve profitability (or reduce operating costs for nonprofit users) can be increased by incorporating quality into the benchmarking analysis. The current chapter proposes an integrated approach to incorporating quality measures into DEA efficiency. The resulting Q-DEA goes beyond the very few DEA studies that have explicitly studied quality and focuses on the importance of the quality dimension.

The quality measure used in the bank application was the one bank management chose to use and continues to rely on. This is only one dimension and one way of measuring quality. One study identifies three dimensions of quality in banking – approachability, location, and telephone service (Athanassopoulos (1997)). Others have identified retention of valuable customers as the key measure. How would the bank application results change if a different quality measure were used? If there were multiple quality measures, as suggested above and as is well recognized in health services, how would Q-DEA or other benchmarking models adapt? This is another exciting area for future research.

This chapter suggests a way of enhancing the DEA methodology for benchmarking and illustrates the potential in a bank setting. While the insights were triggered by the Q-DEA analysis, the motivation to achieve cost savings was strong and other methods might have suggested additional ways to achieve cost savings in the branch network.

Finally, although the discussion is based upon the basic CRS DEA model that assumes constant returns to scale, the quality adjustments considered here could be readily adapted to other DEA models discussed in Chapters 3, 4 and 5.

Chapter 8

APPLYING DEA TO HEALTH CARE ORGANIZATIONS

Hospitals Physicians, Nurses, and Other Caregivers

8.1. INTRODUCTION

Over $2 billion is spent each day on health care in the U.S. - more than 12% of the annual GDP. Costs continue to grow faster than inflation. Health care costs significantly impact individuals through insurance premiums. Federal health insurance programs for the elderly and poor, Medicare and Medicaid claim more and more of citizen's tax dollars. Staggering health care bills, leading to a loss of profits and competitiveness with other countries and between states, burden businesses. Health benefits for retirees originally covered by some large corporations are being reduced or limited in response to the impact on profits. General Motors has declared that there is over $1500 of health care costs in each auto they produce. Even greater pressure on costs is anticipated from the aging population, whose demand for health services is expected to threaten the solvency of the Medicare program. Moreover, the federal government continues to seek ways to cover health care to uninsured citizens, believed to exceed 40 million citizens. This story is paralleled by health care cost inflation in many parts of the world. The problem of rising health care costs, only hinted at above, is well known to most anyone working in the health care field.

Previous cost containment initiatives in health care have been ineffective. In part, this is due to the structure of hospital care, payment systems, and the high cost of new medical technology and pharmaceuticals. Americans hold high expectations about ready access to care and the quality of health care, which hospitals are continually striving to deliver. Insurers have placed few limits on the provision of health services, resulting in negligible incentives to manage the amount and costs of care. Health Maintenance Organizations (HMOs) were established originally as a method of improving quality of care through greater access and a method of cost containment through competitive responses. HMOs are health insurers that receive health care premiums mostly from employers with some employee participation in covering these fees. The HMOs provide a full range of care including visits to primary care physicians (PCPs), specialist physicians, hospital services, therapy, drugs, and to varying degrees other services ranging from eyeglasses to health club memberships to promote preventive care. (A variety of variants on the original health plans have emerged, with names like preferred provider organizations – PPOs. These are also health plans and most of the discussion of plans and HMOs apply to these other forms of health plans.) Unfortunately, HMOs have not reduced the cost of health care (Miller and Luft (1991)).

In 1994 and 1995, several HMOs in several parts of the country charge the same fees as the prior year and/or reduced fees by up to 5%, offering hope of cost containment at last. It is unclear whether that reflects reductions passed on to payers, or whether this was a reaction to the consolidation of HMOs and increased competition. HMOs potentially can manage physicians and patterns of care and reduce health care costs without reducing quality. Still they, like hospitals, currently lack the tools to do so (Chilingerian and Sherman (1994)). Since that brief moment in the mid-90s when the price of health care seemed to be slowing, the costs have continued to rise faster than inflation through 2005 with no clear sign of relief.

In short, productivity improvements and methods to achieve these improvements to reduce health care costs would be very welcome. Indeed, the concept of rationing health care, deciding who can and cannot receive care and treatments for their illnesses, due to limited resources is explicit in some health systems and implicit in others.

While real limitations in resources require rationing, limited resources due to inefficient use of these resources is a less palatable. reason for rationing. Consider how you would react if you were informed that you cannot receive an important treatment until a later time (or possibly never) because every caregiver able to provide that treatment is already fully committed, operating efficiently, and caring for those that most need the care first. Alternatively how would your react if you were told that the treatment is not available *solely* because caregivers are inefficient, taking much longer to provide the treatment than is required? Neither answer is likely to be comforting; however, the former is a matter of public health policy, an important topic beyond the scope of this book. The latter explanation – inefficiency - is due to poor management, a cause that could be eliminated with improved management. If this productivity problem were remedied, more care would be available and/or the cost of care could be reduced. DEA has the potential to help locate the best practice treatment pattern that can be used to help care givers move to this best practice levels. This would be one approach to lower the cost of health care without impacting quality.

In health care, the number of papers and variety of service providers studied using DEA has expanded since the first papers in the early 1980s. These studies have primarily used DEA as a research tool to measure and understand the level of efficiency in hospitals, nursing homes, hospital departments, physicians operating within a hospital and physicians in health plans such as HMOs. Some studies have been conducted with close cooperation of managers in these organizations or with groups of experts to assess the validity of the DEA results. These studies have provided evidence that DEA offers accurate insights about inefficiencies and that these insights may be useful for policymaking and for guiding managers in improving productivity and containing the related components of health care costs.

In contrast to examples cited in banking and government, in chapters 6, 7, and 9, there are no studies we are aware of that document actual cost reductions in health care costs due to the use of DEA. Use of DEA in research has not resulted in reported actions by health care managers. We nevertheless include this chapter for the following reasons:

1. Health care is one of the largest industries in the US and in many other countries.
2. Cost management problems in health care is extremely severe and the nature of health care, with the provision of multiple services with multiple types of resources, many with no market price, makes DEA a potentially powerful management tool to improve productivity and reduce the cost of care.
3. There are studies that have applied DEA to real health care organization data that can serve as a foundation for future efforts to convert the research into practice. Our hope is this chapter can advance this process. We include one specific

application to an HMO and references to other DEA studies to facilitate this objective.

4. Health care is an industry where quality of service is at least as important as the cost of the service. New techniques to incorporate this critical dimension in DEA studies (see chapter 7) will make this more relevant and credible to health care managers, possibly making the use of DEA a more attractive management tool for this industry.
5. DEA is now accessible to managers with DEAFrontier software included in this book, which may attract health care managers to explore ways this may benefit their organization. We believe one barrier to converting DEA into managerial actions is that the numerous other health care studies on DEA were not completed in cooperation with the health care managers and possibly without their fully understanding and agreeing with the way DEA was applied. DEA results will not be convincing if the model is poorly specified or if the data are incomplete and/or unreliable. Participation of management and physicians throughout and application of DEA in health care is needed to develop results that can be used to impact and improve productivity. Now that readers have access to a DEA program that can be readily used without requiring any familiarity with linear programming, we hope that more health care managers and physicians will lead the process of evaluating whether DEA offers insights that can improve productivity providing health services.

There are several papers that have reviewed, summarized and contrasted DEA research studies in health care organizations (see, e.g., Chilingerian and Sherman (2004) and Hollingsworth et al. (1999)). These references can provide examples for managers with studies in health care organizations that supplement the material in the balance of this chapter.

The next section provides an overview of the types of health care organizations that have been analyzed with DEA, DEA models and the input-output data used for these organizations, and comments on findings and potential for future use of DEA to improve operations of these health care units. This section is not designed as a literature review; however, the references in the prior paragraph supply a relatively elaborate guide to DEA health care studies.

Section 8.3 describes a detailed application of DEA to the physicians associated with a large health plan (HMO) and provides detailed review of the types of data used, the results of the DEA analysis, the implications of the results and the management reactions to the results. We believe this is an example of one DEA application that can be used to contain health care costs by identifying and promoting adoption of low cost-high quality physician practice patterns.

The final section 8.4 suggests guidelines and issues that health care managers may want to consider when applying DEA to their organization.

8.2. DEA APPLICATIONS TO HEALTH CARE

DEA studies on health care performance have focused on hospitals, nursing homes, and physicians. Hollingsworth *et al.* (1999) counted 91 DEA studies in health care. The health applications include health districts, HMOs, mental health programs, hospitals, nursing homes, acute physicians, hospital departments, physicians treating patients within hospitals and primary care physicians. These studies include health care organizations in the US, Europe and South America. The DEA input-output models used and the types of insights generated are summarized in the following sections.

8.2.1 Acute Care General Hospitals and Academic Medical Centers

Acute care hospitals have received much research attention using DEA. These studies measured overall technical efficiency of hospitals defining outputs as patient days or patients discharged. These studies tend to focus on the DEA efficiency scores to suggest which types of hospitals are more efficient and to suggest the degree of inefficiency present in individual hospitals and in the hospital system. For example, Hollingsworth *et al.* (1999) found a greater potential for improvement in the U. S. with an average efficiency score of 0.85, and a range of 0.60-0.98, in contrast to Europe with an average efficiency score of 0.91, and a range of 0.88-0.93. This information is not translated into specific differences in the way these hospitals operate and cannot be translated into specific actions by managers to improve efficiency. For research purposes, they do suggest important questions that need to be addressed and they help formulate hypotheses for subsequent studies.

Most of the hospital studies have merely illustrated DEA as a methodology and demonstrated its potential. Researchers have used very different hospital production models making it difficult to compare results. Some combine patient days with patient discharges (essentially counting the number of patients released from the hospital) as outputs, and others separate the manager-controlled production process from the clinical-controlled process (a distinction explained below in conjunction with Figure 8-1).

Some studies have used output measures that are not sensitive enough to generate meaningful results. For example, a study of 22 hospitals in the National Health Service in the United Kingdom used a four output, five input model (Kerr et al. (1999)). The outputs were defined as: (1) Surgical inpatients and visits, and (2) Medical inpatients and visits, (3) Obstetrics/Gynecology patients and visits, and (4) Accidents and Emergency visits. Without knowing the complexity and severity of patients, raw measures of output will lead to distorted results. If Hospital A receives a lower DEA score because Hospital A admits more "fevers of unknown origin," and performs more combined liver-kidney transplants, hip replacements, and coronary by-pass grafts and Hospital B has more tooth extractions, vaginal deliveries (births) without complications, it is an unfair comparison. Studies using DEA or any methodology comparing hospitals with these types of unrefined data cannot generate insights that managers can rely on to develop methods to improve their performance.

Evaluating acute hospitals requires a large and complex DEA model. To generated DEA results that can be relied on and used by managers and policy makers, the models need to capture the complexity of the service unit studied and the inputs and outputs need to reflect the key resources used and services provided. While this prescription may be annoyingly obvious, many studies don't meet this test. The applications to government and banking described in this book were based on DEA models that were subject to management scrutiny and accepted by management as reasonable. This gave the managers the confidence in and the basis for relying on the DEA results. This may be one of the reasons DEA results have not had reported impact on management of health care organizations.

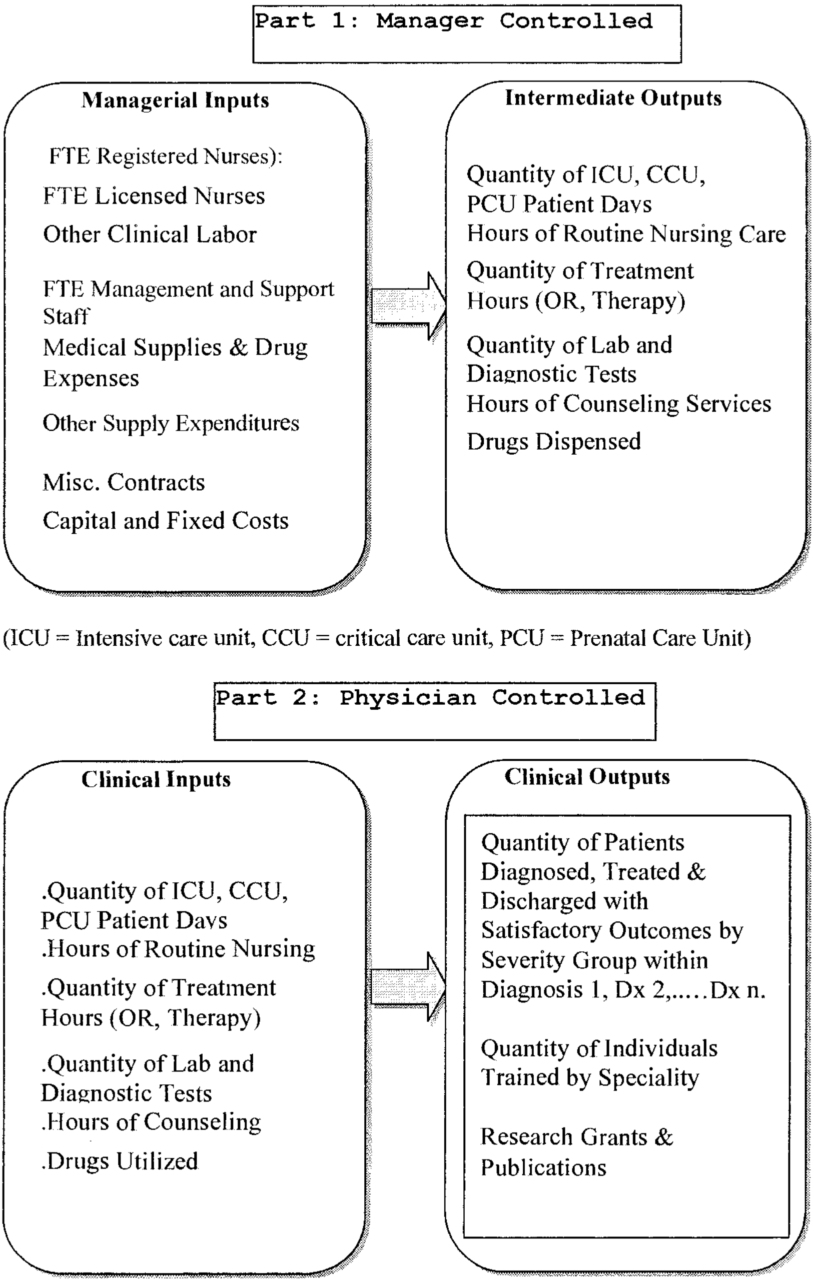


Figure 8-1. Acute Hospital As Two-Part DEA Model (Chilingerian and Sherman 1990)

Nevertheless, there have been innovative hospital-level studies potentially useful for policy makers. For example, dozens of DEA papers have focused on the association between hospital ownership and technical inefficiency studying several thousand hospitals as decision making units (DMUs) (see for example, Burgess and Wilson (1996)). DEA studies have also focused on critical health policy issues such as: regional variations (Perez (1992)), rural hospital closures (Ozcan and Lynch (1992)), urban hospital closures (Lynch and Ozcan (1994)), hospital consolidations (Luke et al. (1995)), and rural hospital performance (Ferrier and Valdmanis (1996))[1].

Examples of possible models for hospital DEA studies found in the literature are presented below. Together, these provide a foundation or starting-point for managers to consider what inputs and outputs need to be incorporated into a hospital DEA study to ensure that the results will be actionable and/or meet other objectives of the analysis. Figure 8-1 describes the medical service production system as a two-part service process: (1) a manager-controlled service unit, and (2) a physician-controlled service unit. In the diagram below, the intermediate outputs of the manager-controlled production process become the clinical inputs for the physician-controlled production process. A discharged patient is the final product, and the clinical inputs are the bundle of intermediate services that the patient received. This model applies to hospitals that include academic medical center (teaching hospital) activities such as research and training of physicians and nurses.

In this model, hospital managers set up and manage the assets of the hospital. They control the labor, the medical supplies, and all expenditures related to nursing care, intensive care, emergency care, and ancillary services (such as lab tests, radiology, and other diagnostic services), pharmacy, dietary, as well as laundry, central supplies, billing, and other back office functions. However, these departments (or functions) merely produce intermediate services that are available for utilization by physicians (see Chilingerian and Glavin

[1] There have, however, been important methodological comparisons. See Banker, Cannon and Strauss (1986) who found increasing or decreasing returns to scale to be present in "each" hospital in contrast to the translog regression which found that constant returns prevailed over "all" hospitals in a North Carolina study. Also see Lawrence Fulton, Performance of Army Medical Department Health Delivery Components, 2001-2003: A Multi Model Approach (Austin, Texas: University of Texas Ph.D. Thesis, 2005, Red McCombs School of Business) compared DEA with SFA (Stochastic Frontier Analysis) and found the former to be superior in identifying sources of inefficiencies. (Also available from University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, Michigan 48106).

(1994), Chilingerian and Sherman (1990) and Fetter and Freeman (1986)). Physician decisions determine how efficiently these assets are utilized. Once a patient is admitted to a hospital, physicians decide on the care program – i.e. the mix of diagnostic services and treatments, as well as the location and intensity of nursing care, and the trajectory of the patient. Physicians decide how and when to utilize nursing care, intensive care, emergency care, ancillary services, and other clinical inputs.

The productive efficiency of the hospital is complicated. A hospital can be clinically efficient, but not managerially efficient. A hospital can be managerially efficient, but not clinically efficient. More often, both parts of the production process are inefficient. If physicians over utilize hospital services ratios such as the cost-per-patient day, cost-per-nursing hour, cost-per-test are reduced, primarily because fixed costs are spread over more units of service, giving the appearance of improved productivity and suggesting that the hospital appears to making the best use of its inputs. However, overuse of these resources just increases the total cost of care and can be a false sign of efficiency.

To be efficient, clinical and non-clinical managers must perform two tasks very well. Clinical managers must manage physicians' decision making (i.e., patient management) and non-clinical managers make the best use of all hospital assets by managing operations (i.e., practice management). Therefore, patient and practice management require an extraordinary amount of coordination and commitment to performance improvement. (In short, as in the simple examples in chapter 2, the hospital needs to reduce the costs of services such as therapies, tests, and bed days - managerial efficiency; and it needs to manage the mix and minimize the volume of these services to minimize the total cost of care while ensuring that the quality of care meets the hospital standard – clinical efficiency.)

Other examples for acute care hospital production models appear below. These are single-stage models and illustrate different ways the inputs and outputs configurations. Figure 8-2 (Burgess and Wilson (1998)) is a model that includes five types of labor inputs and weighted beds as a proxy for capital, but excludes drugs, medical supplies and other operating expenses. Figure 8-3 (Sexton et al. (1989)) collapses nurses into one category, but adds physicians and residents and excludes beds. Figure 8-4 collapses labor into one variable, includes other operating expenses and beds, but also adds a proxy measure of capital based on a count of the number of specialty and diagnostic services.

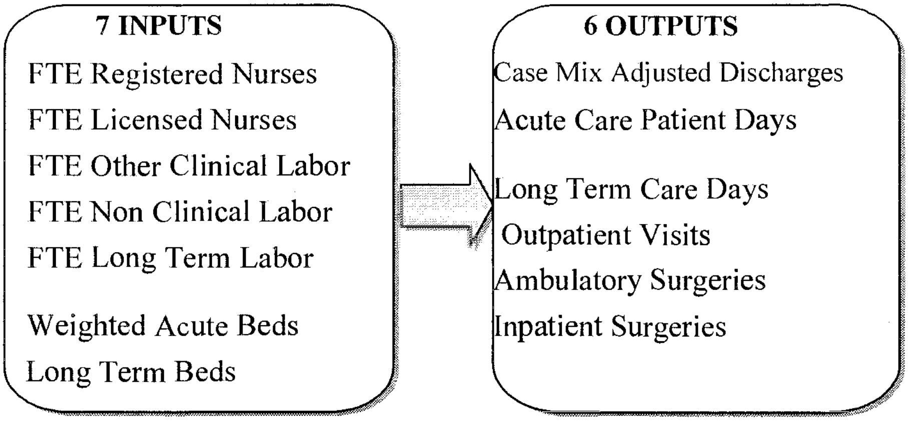


Figure 8-2. Variables in General Acute Hospital Model (Burgess and Wilson 1998)

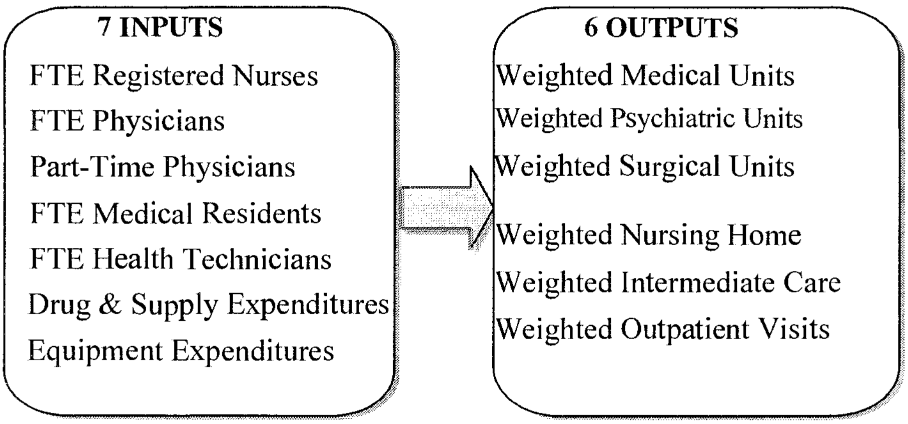


Figure 8-3. Variables in A Medical Center Study (Sexton et al. 1989)

The outputs are different in all three models. Conceptually if the inputs are costs, then the input/output ratios are cost per case, cost per procedure, cost per visit, or cost per nursing day. If the inputs are beds or FTE labor, then the input-output ratio is represented by labor utilized per admission, labor utilized per patient day, labor per surgery, and the like. Mixing managerial inputs with clinical outputs provides valuable insight; however, unlike the two-stage model, managerial and clinical inefficiencies are indistinguishable.

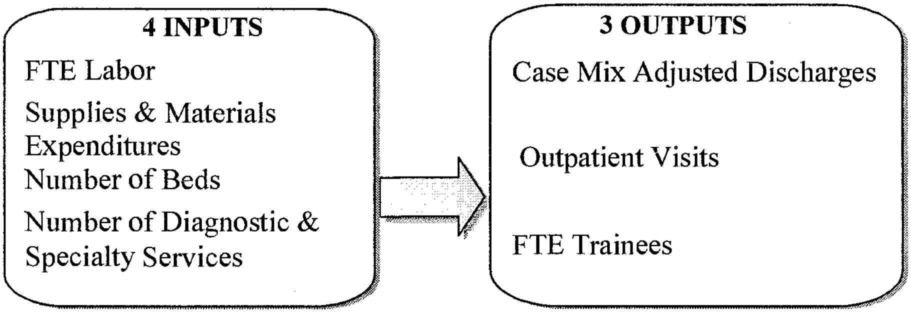


Figure 8-4. Variables in Urban Hospital Model (Ozcan and Luke 1993)

Cautionary note about output measures in health care

Output measures: In this hospital section and in all applications of DEA to health care services, the measure of services needs to consider the severity of the patient illness, the case mix of illnesses treated, and other differences among patients that can require different levels of care requiring different levels of resources. At the extreme, for health care and for other types of services, one might argue every patient and every customer is different with a unique set of service needs. This view would reject the above models along with other methodologies for analyzing health services. More practically, while every patient and customer is different, there are groups of patients similar enough to make comparisons meaningful. This is a judgment that we believe is most appropriately made by the manager that will use the analysis for decision making purposes and who will be responsible for the initiatives based on DEA findings.

Quality measures: Note that none of the above models include quality of care measures. However, no study of health care productivity or costs can be argued to be reliable and comprehensive if quality of care is not considered. Quality is, in itself, a multidimensional issue that needs to be considered. Yet, quality measures are also subject to debate and almost any set of quality measures can be questioned as to whether they are comprehensive and meaningful.

For example, one widely accepted way to gauge quality is to measure the outcome of the treatment, which determines whether the patient has benefited from the treatments provided given their state of health before treatment began. This requires tracking a patient long after they leave the hospital. Only a small few systems track the patient for as much as four months after they leave the hospital. Systems that do not follow outcomes in this way can consider a surgical operation successful even if the patient is readmitted for surgery to correct this same condition at a later point in time.

Another example is the current initiative to encourage hospitals to publicly report on a group of 10 quality measures under the US Medicare system. This was developed by the U.S. Department of Health and Human Services and can be accessed through the website: www.hospitalcompare.hhs.gov. One of these quality measures is the administering aspirin to patients with myocardial infarction (a type of heart attack). The 10 measures relate to only three of the hundreds of types of illnesses treated by a hospital. Discussions with the directors of hospitals suggests that most hospitals will be sure to give every heart attack patient an aspirin along with meeting the other 9 quality measures. While it will be tempting to use these public and accessible quality measures for hospital studies, they can readily be questioned for their completeness and one should not be surprised at the incredibly high percent of hospital that demonstrate outstanding quality of care as measured by these 10 measures. Efforts to increase the number of measures are already in process, and the expanded set may also have similar weaknesses.

Our view is that, like severity and case mix, the quality measures used for health care and other services should be those that the managers believe to be reliable and comprehensive. One test is whether the hospital managers and the health care providers – physicians and nurses – believe the quality measures can be used for analysis and decision making and that they will accept the responsibility for decisions made based on those quality measures.

8.2.2 Nursing Homes

Nursing home studies in the United States often segment the outputs by sources of payments: private insurance, Medicaid, Medicare, private pay. Figure 8-5 displays the inputs and outputs often used in DEA nursing home studies. The nine resources (inputs) are full time equivalent (FTE) registered nurses, FTE licensed practical nurses, and FTE nurse aides, FTE other labor, and medical supplies and drugs, clinical and other supplies, and claimed fixed costs (a proxy for capital). Since DEA can handle incommensurable data, the FTEs are in quantities, and the supplies, drugs and fixed costs are measured by the amount of dollars spent. The outputs are the quantity of nursing home days produced during a given time period. In Figure 8-5, the outputs are the quantity of resident days broken into three payment classification groups: Medicare patients (a national program to pay for elderly care), Medicaid patients (a state program to pay for impoverished residents), and Private patients (residents without financial assistance).

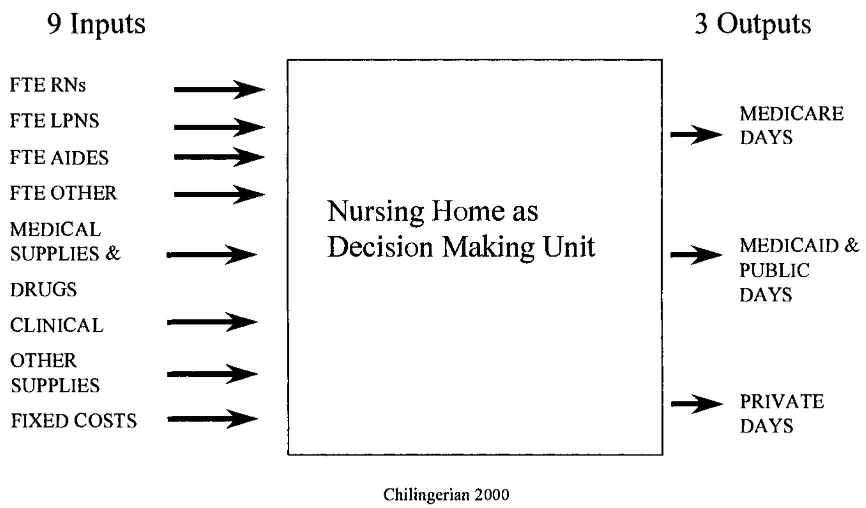


Figure 8-5. Nursing Home Inputs and Outputs (Chilingerian 2000)

This model is not sufficient when the outputs are not homogeneous due to unaccounted for case mix. For example, if the nursing home has skilled nursing which provide care for particular illnesses such as an Alzheimer unit, and/or is a large proportion of patients have some of the following characteristics: older than 85, confused, requiring feeding, bathing and toilet assistance, then the model is not measuring differences in pure productive efficiency. Other measures of case mix for nursing home patients are captured in analysis of the patient's (or resident's) ability to handle activities of daily living such as the ability to walk independently, bathe oneself, and feed oneself. Quality of care is also absent from the model so any results do not recognize whether differences in resources used are due to different levels of service quality. The US Department of Health and Human Services office that oversees Medicare and Medicaid has developed a set of nursing home quality measures that parallels the hospital quality measures and is available on the website: www.medicare.gov/NHCompare/home.asp. The issues related to hospitals apply discussed in the prior section apply to nursing home quality measures as well. Ideally, all nursing home studies should include measure to adjust outputs for severity and for mix of residents.

8.2.3 Primary Care Physician Models: An Example of Clinical Efficiency

With growth of managed care, the primary care physician (PCP) has emerged as an important force in the struggle for efficient and effective medical care. Physician initiate and approve lab and radiology tests, prescription drugs, surgeries, and referrals to specialists and hospitals, essentially defining the amount and type of care provided and impacting the cost of care. Consequently, physician report cards or profiles have become a way to benchmark physician practice patterns with respect to cost of care. Managers can use DEA as a tool to profile and evaluate physicians.

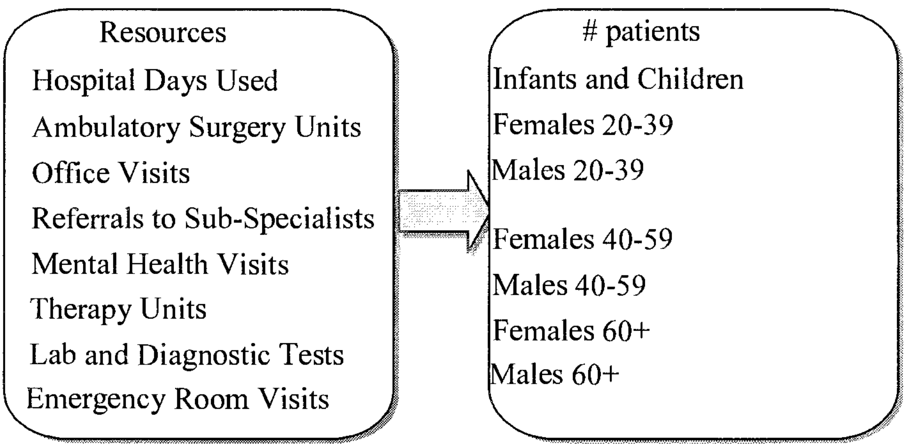


Figure 8-6. Variables in A Primary Care Physician Study (Chilingerian and Sherman 1997)

Previous research has found that three patient variables drive managed care costs. They are: patients' age, gender, and geographic location. Consequently managed care organizations often set their budgets and prices based on these variables. The final product produced by physicians in managed care organizations is one year of comprehensive care for their patients, referred to the physicians panel of patients. To care for patients, primary care physicians (PCPs) utilize office visits, hospital days, lab tests, and therapy units. Figure 8-6 is an example of how one large HMO conceptualized a physician DEA application. Note that quality of care is not included in this model.

8.2.4 Hospital Physician Models: Another Example of Clinical Efficiency

Physicians working within the hospital, whether they are independent or on the staff of the hospital, utilize hospital resources to provide care to a mix of patients. Models for hospital physicians are distinct from the primary care physician working in a plan. For example, a study of 120 cardiac surgeons evaluated how efficiently they performed 30,000 coronary artery by-pass grafts (CABG) on patients over a two-year period (Chilingerian et al. (2002)). Figure 8-7 illustrates a two-input, four-output clinical production model used in that study. The two inputs are defined as (1) the total length of stay (days) for the CABG cases handled, and (2) the total ancillary and other charges (dollars) for the CABG cases handled. The ancillary and other charges input category includes ancillary, drug, equipment, and miscellaneous charges. The first input, length of stay, represents a measure of the duration of CABG admissions and the utilization of clinical inputs such as nursing care and support services. The second input, ancillary and other charges, represents a measure of the intensity of CABG admissions costs and the utilization of operating rooms, laboratory and radiological testing, and drugs.

The four classes of clinical outputs represent completed CABG surgery cases. Since patients with more severe clinical conditions will likely require the use of more clinical inputs, the efficiency analysis must account for variations in case mix in order to be fair to surgeons or hospitals treating relatively sicker CABG patient populations. Accordingly, the outputs are defined by diagnostic category and severity level within diagnosis. In this example, a system of case mix classification called Diagnostic Related Groups (DRG) are used to segment outputs by complexity; moreover, a severity system called MEDSGRPS was used to further segment each DRG into low and high severity categories. The researchers treated DRG 106 and DRG 107 as separate clinical outputs because a CABG procedure with catheterization is more complicated and requires more clinical resources. As explained above, each DRG was further divided into low-severity and high-severity cases.

Studies of physicians within a hospital naturally focus on particular illnesses, which in the above example were two specific DRGs. The model proposed in Figure 8-6 and in the following section is for physicians in a plan or clinic and it considers all patients and all illnesses treated by physicians. These represent one way to model physician practice patterns. Another approach would be to study physicians' treatment of selected illnesses and the resources used to treat patients with those illnesses. Examples of this might be treatment of high volume illnesses such as asthma.

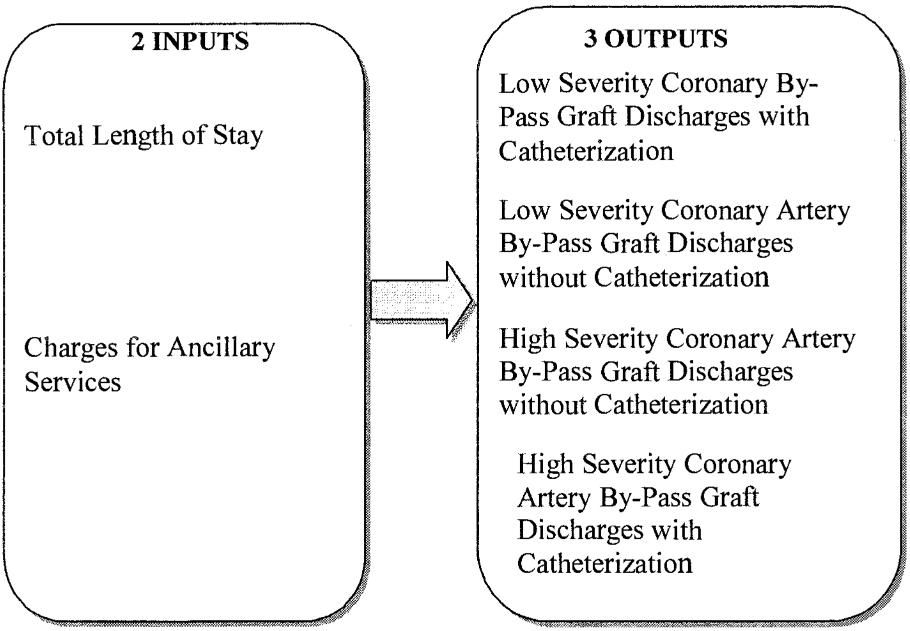


Figure 8-7. Variables in A Cardiac Surgeon Model of DRG 106 and 107 (Chilingerian et al. 2002)

The following section provides a detailed example of applying DEA to plan physicians in cooperation with the medical director and other managers in the plan.

8.3. BENCHMARKING PHYSICIAN PRACTICE PATTERNS WITH DEA: A MULTI-STAGE APPROACH FOR COST CONTAINMENT

This section describes an application of DEA in health care. Physician practice patterns in a Health Maintenance Organization (HMO) are analyzed using single and multi-stage applications of DEA. Best practice (BP) patterns are identified which can serve as benchmark targets for inefficient physicians. We include this example because it suggests the potential value of DEA in health care and was completed in coordination with the HMO's medical officer and his staff. While no specific changes in the operations of the HMO can be claimed from this study, we describe the reaction of the managers and the new insights derived from DEA not available from other analytic techniques used by this HMO.

Three health policy resource utilization control strategies were identified in this DEA study.

1) If managed care organizations could motivate primary care physicians to adopt the practice styles of the best practice primary care physician, substantial clinical resource savings could be achieved, ranging from 12% to over 30% in the HMO that is the focus of this study;

2) Some specialists who practice as primary care physicians (PCPs) provide more efficient care than some general practitioner PCPs, modifying the current perception that reducing specialists is the most effective way to achieve low cost practice patterns; and

3) Groups of physicians in the HMO exhibit different resource use patterns, which may present opportunities to manage high cost groups as another path to contain costs.

The results suggest specific new initiatives, which may prove effective at reducing health care costs within managed care organizations. A multi-stage DEA technique is used to locate specific types of inefficient physicians.

8.3.1 Why Focus on Primary Care Physician (PCP) Resource Utilization?

In this study, we focus on how primary care physicians (PCPs) use clinical resources to treat their patients, rather than fees paid to physician. This focus is driven by the following characteristics of the PCP.

1) Physicians control 80% of health care costs, making this the area where changes can have substantial cost impact (Eisenberg, 1986).

2) Physicians who provide similar quality of health care use distinctly different practice patterns. This suggests there are a number of effective ways to treat similar patients, some of which are less costly than others (Eisenberg and Nicklin (1981); Wennberg and Gittlesohn (1982)).

3) If we can locate efficient practice patterns and help other physicians adopt these patterns, the savings are potentially substantial.

For example, a study of 16 heart specialists associated with one hospital indicated that the more efficient physicians use about 15% fewer resources than the inefficient physicians (Chilingerian and Sherman (1990)).

4) Physician practice pattern management represent an area of growing experimentation and interest for researchers and practitioners (Stein (1993)). Many diverse approaches are being explored. It is unclear whether any of these approaches is sufficient, meaningful, or effective in measuring or improving physician performance.

For example, we conducted interviews at four New England HMOs to determine how they manage costs. Each HMO is pursuing a different approach to managing their physician practice patterns; each is unaware of what the others are doing. These approaches include: establishing effective and economical protocols for specific illnesses; identifying high cost outliers *within the HMO* and working with physicians to reduce costs or removing them from the HMO; and, finally, identifying high cost outliers by comparing physicians with large data bases of physicians throughout the U.S.

The next section describes the structure of this HMO physician study and the use of DEA to accomplish these objectives. Section 8.3.3 reports the DEA results and potential cost savings that might be achieved if all the HMO physicians studied adopt best practice patterns. Section 8.3.4 discusses the results of analyzing specialist versus generalist physician practice patterns. Section 8.3.5 presents insights generated by DEA about physician group practice patterns.

8.3.2 Measurement of Clinical Best Practice Patterns for HMO Physicians

Best practice (BP) physicians are those who use fewer resources relative to other physicians in an HMO when treating a similar mix and number of patients. BP is a benchmarking concept. All physicians within the HMO meet the organization's quality standard. Some BP physicians may exceed this quality standard to a lesser degree than some inefficient physicians. While this is conceptually and theoretically an important issue, HMO managers indicated that they would not retain physicians if they did not meet their quality standard. (Quality measures were not available at this HMO when this study was completed. Including quality measures in physician studies has become more feasible and they are important to ensure that the results don't lead to reducing costs while reducing quality of care below the organization standard.)

Inefficient physicians use more resources than BP physicians to treat a similar mix of patients. Potentially, they could provide the same quality and volume of service with fewer resources.

Identifying BP and inefficient practice patterns using data envelopment analysis (DEA)

Previous resource utilization studies have identified the central tendency or average behavior of physicians. These studies have generally found that the average family physician utilizes fewer lab and diagnostic tests and hospital days than the average general internist or internal medicine sub-specialist. Such studies have relied on multiple regression techniques to identify these central tendencies (Chilingerian and Sherman (1997)). Lewin and Minton (1986) have suggested that multivariate regression models, analysis of variance, and least squares estimation techniques will not lead to adequate explanatory models when the research objective is to identify and analyze best practice. Statistical models based on central tendency theory merely maximize an explanation of average behavior. To find efficient behavior, a technique is needed that can utilize the information in outliers (Lewin and Minton (1986)).

Evaluating a primary care physician's utilization behavior requires an ability to find BPs i.e., the minimum set of inputs to care for a given panel of patients. (Panel is the term used in health care to denote the group of patients assigned to a primary care physician). Since the mathematical relationship between primary care inputs and multiple outputs as defined by the age and gender of patients is unclear, a method of analyzing efficiency and the nature of efficient relationships is needed.

DEA can be used to classify how efficiently physicians use resources. More specifically, DEA measures the magnitude of departure from the BP frontier for each physician based on his or her use of primary care resources to care for a panel of patients. DEA separates all physicians who define the BP frontier from inefficient physicians lying off the frontier by analyzing the multiple resources used (inputs) to provide patient care (outputs).

DEA accomplishes four analytic tasks. First, it classifies physicians into peer groups of physicians who are similar in their use of inputs to produce outputs. Second, it defines a BP (relatively efficient) frontier consisting of all physicians providing their mix and volume of care with the least amount of resources (inputs). Third, it creates an index which measures how far each inefficient physician is from the BP frontier - the DEA efficiency rating. Fourth, it converts the efficiency rating into an estimate of the excess resources used by each inefficient physician compared with BP physicians - the potential cost savings.

In this study, we define the decision making unit that will be evaluated as the individual primary care physician (PCP) -- a term used widely by HMOs. Some PCPs have subspecialties but treat the same types of patients as the generalist PCPs. To construct the BP production frontier, observed behavior is evaluated by using the following input-output criteria:

a) A PCP is inefficient if he or she could provide the same volume and mix of patient care at or above the HMOs' quality standard with fewer input-resources.

b) A PCP is inefficient if it is possible to increase the amount of patient care he or she provides at or above the HMOs' quality standard without decreasing the amount of input-resources used.

A PCP is classified as efficient only when both criteria indicate that the physician is not inefficient (see chapter 2). The BP physicians may be characterized as follows: based on comparison with other physicians in the study, "it is not possible to improve some observed input or output value" for the best practice physician "without worsening other input or output values." (from Banker et al. (1989), page 141).

This definition of BP physician does not necessarily coincide with the lowest cost, high quality practice pattern, as we are focusing on technical efficiency rather than allocative or price efficiency. We are identifying ways to reduce resources that will ultimately lower cost. However, a reduction of one unit of a resource, such as hospital bed days, can shrink costs more than eliminating one unit of another resource, such as tests or therapy treatment. Allocative efficiency focuses on the lowest cost mix of resources If we knew the cost of each input used, we could include that information to determine which were the lower cost patterns, and thereby incorporate allocative and technical efficiency into the same analysis. Unfortunately, the costs of the inputs are not available on a per unit basis and are not all purchased on a per unit basis. Future studies where such information is available could locate a more refined set of best practice physician patterns that will be technically and allocatively efficient. This could help management adopt an even lower cost set of practice patterns than we seek in this study.

For the following reasons, we utilize the CRS DEA model, which assumes constant returns to scale, and not the variable returns models, which filter out returns to scale effects. In this study, the PCPs annually treat from 50 to more than 200 patients that are member of the HMO. These physicians may each have 2000 or more (some as many as 5000) patients, which include those from other HMOs, insurance plans, and private patients. Hence, we don't have data on the entire patient panel of each physician, where returns to scale may be real and important in terms of cost and quality. More importantly, we assume that two patients with similar illnesses will need two times as much care as one patient. No evidence exists that having more patients of similar type will reduce the number of days in the hospital, tests, office visits, etc. Health care efficiency studies that build on this could focus on the implications of measuring and interpreting scale effects. This would be particularly appropriate in studies of the entire patient panel for the physicians or a study of services provided by a hospital or laboratory

The study population

We conducted this study at Alpha Health Plan (disguised name), a large Independent Practice Association (IPA) with a network of 3,000 physicians located on the east coast of the United States. Unlike a pure IPA where the physician contracts with the HMO, Alpha contracts with an individual provider unit – a group of physicians in a geographic area. Each group is given a per panel member per month budget. If, at year's end, the HMO is under its budget, part of the budget-surplus is shared with the groups.

This study's data include physician-level observations of every prepaid primary care service provided by PCPs to members (except telephone calls and prescription drugs) for a twelve-month period. The physicians in the study population have all been PCPs in the Alpha Health Plan for at least one full year with average panel sizes of 50 or more. (The panel refers to the group of patients assigned by this HMO to a primary care physician, who may be a small proportion of the entire group of patients treated by that physician).

We limited the current study population to three primary care specializations: generalists in family practice, and specialists who are general internists with and without board-certified sub-specialties. After applying the criterion, the study population included 326 physicians: 86 in family and general practice with an average panel size of 281 patients (members), 169 general internists with an average panel size of 217 patients, and 71 board-certified sub-specialists with an average panel size of 175 patients. (The sub-specialist physicians are all general internists with sub-specialties in 13 areas such as cardiology, oncology, and dermatology. No one sub-specialty included more that 15 of these 71 physicians). In total, these physicians are responsible for providing care to 73,000 patients that are members of this HMO plan.

PCPs that are internists or internists with sub-specialties are employed by the HMO to serve as the PCP and not as a specialist. Generally, the specialists that a PCP refers a patient to in this HMO will not be another PCP with a sub-specialty. In addition, the specialist PCPs are not paid more because of the sub-specialty and the groups receive the same per-member per-month budget regardless of the number of specialist-PCPs in their group. Hence the specialists and generalists are treated the same and face the same production frontier. There is some possible self-selection among patients that may cause a patient to prefer one physician to another. This may cause some patients to seek a PCP with a sub-specialty just as it may cause a patient to seek a physician who has a friendly caring manner. This HMO had no data that would suggest whether this selection process skewed the patient types among the different categories of PCPs.

Measures of resource use

We obtained measures of physician resource use by collecting utilization and cost data for each patient in the PCP's patient panel for one full year. Using the concept of the PCP as the gatekeeper controlling all care to HMO members, if a PCP referred a patient to a specialist, the referring PCP is accountable. for the referral visit. The

PCP is responsible for providing care to his or her patient panel and has no control over the types of illnesses that will require treatment each year. An efficient PCP will use fewer resources than an inefficient PCP to provide care to their patients. The model in Figure 8-8 and the DEA application in this study follow this view of the physicians' activities and areas of control.

The resource used were: medical/surgical office visits with the PCP; ambulatory surgery procedures; medical /surgical hospital days; mental health office visits; therapy visits (i.e., physical therapy, occupational therapy and speech); radiology, lab and diagnostic tests; emergency room visits, and referral office visits with a specialist. These resource categories reflected the HMO's view of the resources that PCPs control and are responsible for. Their information system could readily provide this type of data.

Measures of case mix

The mix of patients in a panel (sometimes referred to as the PCP's case mix) influences the utilization of clinical and technical resources. According to the literature, the utilization of primary care resources is viewed as an artifact of the various medical care experiences of the panel populations. Studies have demonstrated that risk profiles of a panel can be captured (by proxy) by the age and gender mix. Studies have found that three variables drive HMO costs: age, gender and geographic location; consequently, most HMOs set their prices based on these variables (InterStudy 1989). Age and gender were used in this study to control for case mix. Figure 8-8 illustrates the model of inputs (resources) and outputs (patients treated) used in this study.

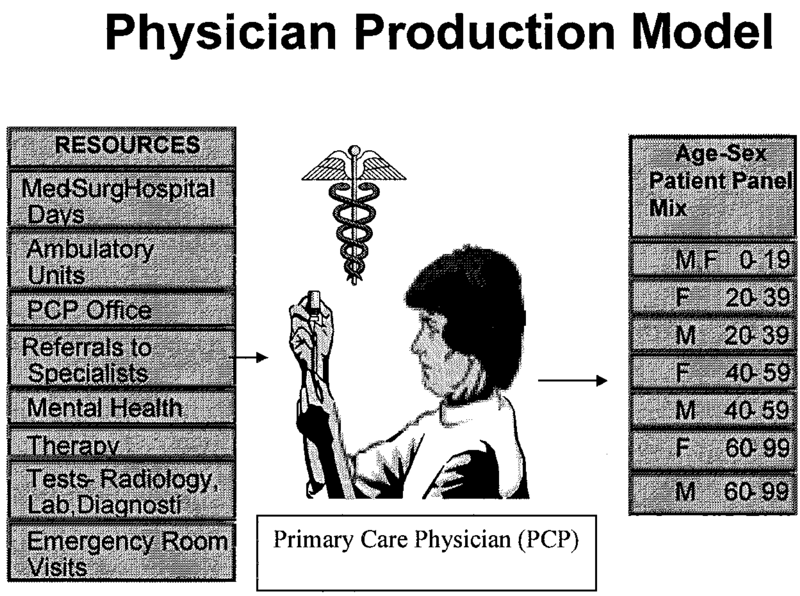


Figure 8-8. DEA Model for Evaluating Efficiency of Primary Care Physicians (PCPs)

In Figure 8-8, Resources used by each PCP included in this study are listed on the left. The resources are used to treat the number of patients treated by each PCP in each age sex group listed on the right. A best practice PCP would use the minimum amount of resources to treat their panel of patients while meeting the health plan quality standard.

Case mix and quality - qualification

Notwithstanding the justifications for using age-sex criteria to control for case mix, health status and severity dimensions may require more sensitive case mix measures to evaluate physicians fairly. Interviews at the four HMOs we studied indicated a lack of sensitive case mix data, limiting their ability to manage physician costs, regardless of available analytic techniques. HMO managers will inevitably be forced to address this issue as cost-based competition becomes more pronounced. Improved tracking of case mix data is critical both for management control, and for negotiating contract fees that ensure health services generate revenues that exceed costs (Stein (1993)).

Similar limitations exist when measuring the quality of care physicians provide. While some surrogates measures are available, such as patients leaving a physician and malpractice suits, direct measures are only beginning to be developed (Stein (1993)). When asked to identify low quality physicians, HMO directors have replied that they would not retain physicians who did not meet their organization's quality standard. In 1999, the largest US HMO announced that if would begin to measure quality of care provided by their physicians for the first time. Today, HMOs have increasingly been capturing quality of care via tracking the frequency in which plan members choose to change physicians, customer satisfaction surveys and measuring the percentage of their panel that receive basic services such as immunizations, mammography, annual checkups, and standard diagnostic tests. To develop initiatives that can reduce costs and maintain or improve quality, quality measures need to be included in productivity analyses. Some of the ways this can be accomplished including a method referred to as quality adjusted DEA (Q-DEA) are described in chapter 7.

In this application, we define best practice as a physician that uses resources efficiently and we do not include an explicit measure of quality of care. The HMO did not have quality measures at the time of this study.

It is important to note that high quality does not necessarily require more resources and there are times when less resource use results in higher quality care because it may result in fewer invasive procedures that may have undesirable side effects. At the same time, we believe that there are differences in quality of care among PCPs even if the HMO manager does not measure or acknowledge them. Hence there may be physicians that are somewhat inefficient in use of resources but who have such superior quality that one would consider them true models of best practice. Similarly, there may be two relatively efficient BP physicians that would be rated equal by DEA in terms of resource use. If one has much higher quality of care, the plan managers might treat the high quality doctor as a model and not the other lower quality physician as a model, even though both meet the minimum quality standard. This type of analysis will be possible when reliable and accurate managed care health plan quality measures are more readily available.

8.3.3 Research Findings - Potential Cost Savings Classification of Physicians into BP and Inefficient Practice Styles

Working with the medical director of Alpha, we developed the seven output-eight input DEA model in Figure 8-8 to classify the 326 primary care physicians into the practice style categories: BP (on the relatively efficient frontier) and Inefficient (off the relatively efficient frontier) based on their observed utilization behavior. (Details on inputs and outputs are in Appendix A). The eight inputs were deemed discretionary inputs by the Medical Director, including emergency room visits and therapy visits. The relative performance of each of the 326 physicians was calculated in relation to all others, based on observed values for the inputs and outputs. One hundred and thirty eight were practicing on the BP frontier and 188 were inefficient with DEA rating of less than 1. (See Table 8-1).

The estimated degree to which the inefficient physicians used excess resources and their distance from the BP frontier suggests wide variations in physician practice behavior. The distribution of DEA ratings of the PCPs is reported in Table 8-2. Three PCPs had ratings below 0.500. If these doctors adopted BP patterns, they could provide the same quantity and quality of care with about one-half of the resources they used during the study year. One hundred and fifty-two physicians had DEA ratings below 0.900 indicating that they have potential to reduce the resources they use by at least 10%. Based on other DEA applications, HMO managers should focus on lower score physicians (below 0.900) first to locate the PCPs able to reduce resource utilization. Data problems and input and output specification problems (which management may not be aware of) can cause the results to be skewed causing some of these PCPs with scores between 0.9 and .99 to have mathematical inefficiencies due to minor data problems rather than real operating inefficiencies . While the data problems may explain and justify one or more physicians using 10% more resource than BP physicians, they are unlikely to justify use of more than 10% of excess resource for most of the physicians with lower DEA ratings. (The cutoff of 10% is somewhat arbitrary, but it has proved practical in other DEA applications, as the higher inefficiencies tend to be real and lead to initiatives that generate real cost savings. When analyzing the results, this cutoff should be scrutinized by managers involved in the process, as even a 2% level of inefficiency can be real and substantial it relates to a large volume of patients.)

Table 8-1. DEA Distribution of Best Practice and Inefficient HMO Physicians

	Generalist	Specialist	total
Best Practice On the Frontier	39	99	138
Inefficient Off the Frontier	47	141	188
total	86	240	326

Resource saving implications for the HMO

We summarize estimates of Alpha's potential savings if all its inefficient physicians adopted BP patterns in Table 8-3. For example, the number of referral visits would be reduced by 14,310, which is 13.7% of the total referral visits for all patients in the study year. Therapy treatments would be lowered by 11,777, which is more than 32% of the yearly therapy visits. The range of resource reductions is lowest in PCP visits (12.5%) and highest in therapy (32.3%).

Such reductions (except for PCP visits) could result in substantial cost savings to Alpha. This magnitude of resource savings would be considered significant by almost any business and this HMO might use this opportunity in several ways. Alpha could, for example, hold premiums steady and increase profitability, or reduce premiums somewhat without reducing profitability.

Table 8-2. Distribution of DEA Physician Ratings

DEA Efficiency Rating Inefficient	# of Physicians	% in each rating level	Cumulative # At or Below the DEA Rating	Cumulative % At or Below the DEA Rating
.400 - .499	3	1.6%	3	1.6%
.500 - .599	5	2.7%	8	4.3%
.600 - .699	22	11.7%	30	16.0%
.700 - 700	45	23.9%	75	39.9%
.800 - .899	77	41.0%	152	80.9%
.900 - .999	36	19.1%	188	100.0%
Total	188	100.0%		
Best Practice				
1.000	138			
	326	Total Physicians in Study		

Table 8-3. Potential Resource Savings if Inefficient Physicians Adopt Best Practice

Type of Resource	Health Care Resources Used by 362 (PCPs) Primary Care Physicians	Potential Resource Savings if all PCPs Adopt Best Practice Patterns	Potential Savings of all Resources Used
Medical-Surgical Hospital Days	26,321	5,785	22.0%
Ambulatory Surgery Units	2,646	616	23.3%
PCP Office Visits	129,611	16,227	12.5%
Referrals Resulting in a Specialist Visit	104,237	14,310	13.7%
Mental Health Visits	44,340	9,209	20.8%
Therapy Units	36,456	11,777	32.3%
Radiology, Lab and Diagnostic Tests	550,598	116,036	21.1%
Emergency Room Visits	11,972	2,170	18.1%

The insights provided by the DEA results at the individual PCP level are illustrated in the example in Table 8-4. This example is indicative of the type of information that can be derived from DEA for each less productive PCP. PCP G234, a generalist, had a DEA rating of 0.580, indicating s/he was an inefficient PCP compared with the other PCPs in the study. (G# is a generalist PCP and S# is a specialist PCP). PCP G234 has potential resource savings of 99 medical-surgical inpatient hospital days, 3 ambulatory surgery incidents, 285 PCP visits, 105 referrals, 38 mental health visits, 35 therapies, 682 tests, and 49 emergency room visits. These potential efficiency gains are based on the actual resources and services provided by the BP physicians (on the frontier) closest to the input-output mix of PCP G234.

The BP Reference Set of physicians for PCP G234 is listed in Table 8-4 and include PCP #'s G150, G153, G154, G197, G339, G87, S340, and S78 (6 generalists and 2 specialists). The weighted sum of the BP physicians suggest a practice pattern that would provide as much or more patient care with fewer resources – referred to as the *best practice target* (see Table 8-4)[2].

[2] The target column amounts are rounded to the nearest whole unit in Table 8-4 and Table 8-7.

Table 8-4. General practice physicians identified as inefficient compared with best practice physicians (BPP)

PCP 234 - Generalist

DEA Efficiency Rating = 0.580

Age-Sex Patient Panel Mix	G234 Actual	Best Practice Target	Savings	Best Practice Reference Set of Physicians Actual Panel and Resources Used G150	G153	G154	G197	G339	G87	340	78
MF 0-19	11	27		32	5	15	65	6	21	7	1
F 20-39	67	67		46	21	36	40	58	371	9	47
M 20-39	55	55		86	46	77	38	64	49	5	19
F 40-59	26	31		17	9	32	43	17	94	8	22
M 40-59	31	31		26	12	29	55	31	23	2	13
F 60-99	3	7		2	4	17	5	1	27	[illegible]	2
M 60-99	5	5		3	9	10	8	1	1	[illegible]	1
Total Patients	198	223		212	106	216	254	178	586	24	104
Resources											
Medical-Surgical Hospital Days	156	56	99	68	46	9	39	42	230	8	95
Ambulatory Surgery Units	8	5	3	5	3	5	0	11	15	[illegible]	8
PCP Office Visits	709	424	285	254	216	175	751	357	726	66	216
Referrals Resulting in a Specialist Visit	261	156	105	112	67	174	69	101	789	83	114
Mental Health Visits	94	56	38	41	20	18	6	78	372	4	16
Therapy Units	70	35	35	13	26	78	1	1	246	[illegible]	7
Radiology, Lab and Diagnostic Tests	1695	1012	682	1211	288	577	1112	1193	3050	79	462
Emergency Room Visits	69	20	49	43	18	17	7	25	47	8	6
DEA LAMBDA COEFFICIENTS				0.089	0.120	0.130	0.277	0.123	0.078	.091	0.049

The weights related to each BP reference physician are from the dual linear program results calculated by DEA as it tries to maximize each PCP's efficiency rating. These weights, referred to as the Lambda coefficients, are reported at the bottom of Table 8-4. (The potential savings reported in Tables 8-3 and 8-4 (and Table 8-7) represent the radial contraction of the resources of the inefficient PCP by the DEA efficiency rating and the elimination of input slacks.

We are focusing on the potential resource savings that are identified with DEA by locating PCPs whose combined (or weighted) practice patterns produce at least the same volume and mix of services as G234 but with fewer resources. At this juncture, the focus is on technical efficiency rather than allocative efficiency, which could lead to added resource savings. Moreover, one could argue the even if G 234 reduced resources by the amounts suggested in Table 8-4, that there would remain some level of inefficiency, as the BP target produces more with the resulting level of inputs. For example, the BP target treated 16 more MF 0-19 (male and female patients age 0 to 19 years old) than G 234 (11 for G234 versus 27 for the BP target) We assume that the patients treated are relatively less controllable and that increasing the productivity of G234 beyond the resource savings by increasing outputs is an issue that should be addressed in subsequent studies and in the context of what is feasible and controllable in an HMO.

The changes in resource use reflected in Table 8-4 would make G234 as efficient as the BP physicians except for the potential to increase outputs. The key point with respect to the objective of this study is that these resource savings alone represent substantial and specific potential cost savings. The issue that remains in this context is how can the HMO realize these potential cost savings. This requires finding ways to help physicians like G234 adopt lower cost practice patterns such as those in the BP reference set. (One alternative is to reduce the number of inefficient physicians, the focus of the next section of this paper)

To analyze the differences in practice patterns that cause G234 to use excess resources, the HMO could compare and analyze the BP Reference Set physicians' practice patterns. These PCPs continued membership in the plan implies that they are providing care at or above the HMO's quality standards. In practice, G234 may be unable to reduce all resources by the precise amount specified by DEA. However, the extent of the possible reductions for the specific patients treated could be determined by more detailed comparison of the practice pattern of G234 with the BP Reference Set of physicians. Based on applications to other types of DMUs (e.g. bank branches discussed in the chapters 6 and 7), the actual savings could be substantial. While the actual savings may be below the potential level for some resources, they also may exceed the potential for other resources. (Similar information on potential savings for each of the 188 inefficient PCPs by type of resource is generated by DEA). The process of developing ways to interpret these practice pattern differences and motivate less productive physicians to help identify and adopt lower cost practice patterns is the subject of a study that is currently in process and an issue readers in health care sector may want to explore.

The results suggest that there are substantial potential savings if the identified inefficient PCPs adopt BP patterns. These cost savings could limit insurance premium increases, provide incentives to physicians who adopt the BP patterns, and reward the BP physicians.

The HMO could use these results for "economic credentialing" of physicians to evaluate their profitability or for creating a "financial" report card. Such analyses could help determine whether the costs of a PCP's practice pattern the fees earned from his or her panel of patients. It could also identify physicians who might be removed from the plan due to excess resource use, or who should be retrained to adopt lower cost practice patterns. Such practice pattern data could help determine whether a physician who management is considering hiring will be accepted to the plan based on prior treatment history.

The BP physicians identified with DEA may provide insights to help improve the HMO practice patterns. The protocols developed by HMOs to guide their PCPs in effective, low cost treatments could be based on the patterns used by the BP physicians. These PCPs could also help improve the inefficient PCPs' practice patterns. The benefits of utilizing the DEA results for these purposes need to be clinically tested before the value of this approach can be measured. The substantial cost savings realized banking and in government services (discussed in the chapter 9) (Sherman, 1988) suggests the value of testing this approach. The results already suggest strategies and initiatives to reduce operating costs without reducing quality or volume of care.

8.3.4 Research Findings: Managing the Mix of Generalist versus Specialist PCP Practice Pattern to Reduce Costs

The study of the Alpha HMO suggests ways DEA can be used to locate specific types of inefficient units, providing new options for the management to achieve improved productivity and reduced operating costs.

Substantial evidence exists that specialist PCPs (such as cardiologists) use more resources than generalists PCPs (general practitioners and family medicine) (Bergman and Pantel (1986); Fishbane and Starfield (1981); Cherkin et al. (1987); Childs and Hunter (1972); Eisenberg and Nicklin (1981); Freeborn et al. (1972); Bertakis and Robbins (1987)). For example, Noren (1980) found that internists tended to use more laboratory and X-ray tests than did family-general practitioners. A medical outcomes study conducted at New England Medical Center revealed that (after controlling for patient mix) sub specialists tended to use more primary care resources than internists, and internists used more resources than family physicians (Greenfield et al. (1992)). Upon release of the study in March 1992, the news media described the rising cost issue in terms of "America's reliance on specialists and the expensive brand of medicine they practice" (Knox (1992)). An editorial in the *Journal of the American Medical Association* argued, "one way to gain some control over escalating health care expenditures is to pay attention to the mix of physicians providing health care...." (Rosenblatt (1992), page 1666).

Mean utilization rates of clinical and technical resources for the three specialty groups are presented in Table 8-5. By and large the results support previous work. Although generalists utilize more office visits and more emergency room visits, they use significantly fewer lab, radiology, diagnostic tests, and med-surg days and make fewer referrals than the specialists. These results argue that HMOs should minimize the employment of specialists as PCPs. One resource that is used more by generalists – family physician visits – would require no added out-of-pocket costs to this HMO. Their only obligation is to pay the PCP based on the total patients in the panel, regardless of number of office visits. Other resources like referrals and tests require added out-of-pocket costs – the types of resource used more intensely by specialist PCPs.

Some HMOs have actively tried to reduce the number of specialists so as to lower their costs. Chilingerian and Sherman (1997) have found evidence suggesting that some specialists can be as cost effective as generalists and some specialists are more efficient than other specialists. This is not a new concept. While only anecdotal evidence supports this reasoning, generalists acknowledge that some orthopedists prescribe surgery more often than other orthopedists for similar conditions.

Patients who will not be harmed but will not benefit from surgery are better off going to the "more efficient" specialist who prescribes surgery less frequently. The latter is also likely to treat the patient at lower cost as a result of avoiding surgery and hospital days. The generalist will often chose the specialist referral based on their assessment of the type of treatment they think will be most appropriate.

Table 8-5. Mean Resource Utilization Rates: Generalists versus Specialists

MEAN UTILIZATION RATES PER THOUSAND MEMBERS	PHYSICIAN TYPE: GENERALISTS	SPECIALISTS: INTERNISTS	SPECIALISTS: SUB-SPECIALISTS
	86 PCPs	169 PCPs	71 PCPs
OFFICE VISITS	1958	1752	1449
MED-SURG HOSPITAL DAYS	242	280	364
AMBULATORY SURG. UNITS	33	37	44
EMERGENCY ROOM VISITS	177	164	180
LABORATORY TESTS	4506	5679	5583
DIAGNOSTIC TESTS	903	1152	1309
RADIOLOGY TESTS	1270	1496	1619
MENTAL HEALTH VISITS	545	646	537
THERAPY UNITS	415	517	460
REFERRAL VISITS	1237	1501	1667

Our hypothesis is that two types of specialists exist - higher cost and lower cost. If true, HMOs can manage costs more effectively by pursuing a strategy of reducing inefficient specialists *and* inefficient generalists instead of reducing specialists because they use more clinical resources, on average.

To illustrate this possibility, a hypothetical set of 13 PCPs in an HMO is presented in Table 8-6. Three are efficient or best practice generalist - BP (D1, D2, and D3), four are generalist - inefficient (D4, D5, D6, and D7), three are specialist - BP (D8, D9, and D10), and three are specialist - inefficient (D11, D12, and D13). The average cost of care per patient (column A, Table 8-6) reflects the general expectation that inefficient physicians will use more resources, resulting in higher cost per patient. This set of physicians was constructed so that some BP specialists are lower cost than some inefficient generalists. For example, specialist D8 has an average cost of $1,500 per patient versus $2,600 per patient for the inefficient - generalist D7. The general picture is consistent with the view that specialists are on average more costly: the average cost per patient for generalists is $1,757 versus an average cost for specialists of $2,317. The average cost per patient for all the physicians is $2,015.

Table 8-6. Impact of Alternative Strategies on HMO Cost of Patient Care

Retaining all physicians versus retaining only generalists versus retaining only best practice specialists and generalists

PCP #	PHYSICIAN DEA CLASSIFICATION ECONOMIC CREDENTIAL	A ALL PHYSICIANS	B RETAIN BEST PRACTICE PHYSICIANS	C RETAIN GENERALIST PHYSICIANS
		COST PER PATIENT		
D 1	Generalist - Best Practice	$1,000	$1,000	$1,000
D 2	Generalist - Best Practice	1,200	1,200	1,200
D 3	Generalist - Best Practice	1,400	1,400	1,400
D 4	Generalist - Inefficient	1,500		1,500
D 5	Generalist - Inefficient	2,200		2,200
D 6	Generalist - Inefficient	2,400		2,400
D 7	Generalist - Inefficient	2,600		2,600
D 8	Specialist - Best Practice	1,500	1,500	
D 9	Specialist - Best Practice	1,700	1,700	
D 10	Specialist - Best Practice	2,000	2,000	
D 11	Specialist - Inefficient	2,500		
D 12	Specialist - Inefficient	2,900		
D 13	Specialist - Inefficient	3,300		
AVERAGE COST PER PATIENT				
	ALL PHYSICIANS	**$2,015**		
	SPECIALISTS	**$2,317**		
	GENERALISTS	**$1,757**		**$1,757**
	BEST PRACTICE	**$1,467**	**$1,467**	

The simple approach of eliminating specialists and retaining the generalists results in an average cost per patient of $1,757, which is lower than the overall average patient cost of $2,015. (see column C of Table 8-6). If the HMO could identify the BP physicians whose practice patterns meet quality standards with relatively low cost of care, it could adopt an alternative policy of retaining these physicians.

Such a policy would include efficient specialists *and* generalists as in this hypothetical example. When all inefficient physicians are eliminated and BP physicians are retained, the average cost per patient would be reduced to $1,467 (see column B of Table 8-6). This strategy produces lower costs than the prevalent strategy of simply eliminating specialists. This is possible only if some specialists cost less than generalists. This strategy is feasible if the HMO can identify and retain the lower cost specialists, and if it can identify and eliminate or retrain the high cost generalists. We suggest below an approach to identify the higher cost generalists using a multi-stage application of DEA.

Multi-stage DEA applications to locate generalists using more resources than specialists

To explore the question of whether some specialists are less costly than generalists, we analyze the resource utilization of these two types of PCPs in this HMO. We introduce the use of a multi-stage application of DEA as a tool to determine if there are generalists that are higher cost than the specialists, and to identify these PCPs. This relies on the ability of DEA to indicate the BP Reference Set of Physicians associated with each inefficient physician. While DEA identifies the inefficient physicians by comparing each physician with every other physician in the study, it also indicates the specific subset of physicians that most directly point to their use of excess resources. In the example of PCP G234 (Table 8-4), the BP reference physicians included two specialists and six generalists. PCP G234 was identified as inefficient compared to a combination of generalists and specialists physicians. This means that some specialists were more efficient than PCP G234 on some dimensions, while some generalists were more efficient on other dimensions. Hence, the practice patterns of specialists only partly identify this generalist as inefficient. Based on this first stage of the DEA analysis, we do not know if G234 used more resources than specialists on the BP frontier.

To focus on generalist versus specialist practice patterns, DEA was run a second and third time comparing the 47 inefficient generalists (off the frontier) with the 99 BP specialists (on the frontier). The objective was to pit these inefficient generalists against the specialists on the BP frontier to address the question: Are any generalists inefficient compared only with specialists and who were these generalists? These inefficient generalists would be distinguished by the fact that they have only specialists in their BP Reference Set.

The multi-stage process proceeded as follows. After the initial results reported in Table 8-1 were received, we reviewed the inefficient generalists and found none had a BP Reference Set with only specialists PCPs. We then reran the DEA comparison of the 47 inefficient generalist PCPs and the 99 BP specialists (in Table 8-1). The results of this second DEA run were analyzed and the 47 inefficient generalists still had a mixture of specialists and generalists in their BP Reference Set. Twenty-two of the generalists now become BPs. This can occur because we removed the 39 BPs that were more efficient and highlighted these 22 physicians as inefficient in the initial analysis (see Table 8-1). The absence of those 39 PCPs as a basis for comparison allows the 22 generalists to appear relatively efficient. A third DEA run was completed removing the 22 generalists identified as BPs in the second run. The third DEA run compared the remaining 25 generalists that were rated as inefficient with the original 99 specialists BPs.

This third stage of the DEA analysis identified eight generalists as inefficient with *only* specialists in their BP Reference Sets. This suggests that at least these eight generalists use more resources than some specialists. This multi-stage DEA application is summarized in the following Table.

	# of GENERALIST PCPs		# of SPECIALIST PCPs		Total
STAGE	BP (best practice or efficient)	Inefficient	BP	Inefficient	
1	39 (removed from stage 2)	47	99	141	326
2	22 (removed for stage 3)	25 None have only BP specialist in their Reference Set	99	NA	146
3	7	18 8 have only BP specialist in their Reference Set	99		124

Table 8-7. General Practice Physicians Identified as Inefficient Compared with Best Practice Specialists

PCP 234 - GENERALIST DEA EFFICIENCY RATING = .688				ALL SPECIALISTS					
Age-Sex		BEST	POTENTIAL	Best Practice Reference Set of Physicians					
patient panel mix	G234	PRACTICE	RESOURCE	Actual Panel and Resources Used					
	Actual	TARGET	SAVINGS	S340	S17	S78	S139	S127	S245
MF 0-19	11	26		17	113	1	11	5	5
F 20-39	67	67		69	92	47	87	23	27
M 20-39	55	74		95	122	19	27	31	39
F 40-59	26	29		18	76	22	33	15	26
M 40-59	31	31		22	95	13	15	21	38
F 60-99	3	4		2	13	2	7	5	8
M 60-99	5	5		2	16	1	7	6	13
Total Patients	198			224	526	104	186	105	155
Resources used									
Medical-Surgical Hospital Days	156	75	81	58	104	95	59	118	107
Ambulatory Surgery Units	8	6	2	5	9	8	8	4	3
PCP Office Visits	709	447	262	366	1268	216	309	117	236
Referrals Resulting in a Specialist Visit	261	200	61	183	221	114	284	142	423
Mental Health Visits	94	72	22	74	160	16	56	22	49
Therapy Units	70	25	45	6	33	7	90	21	90
Radiology, Lab and Diagnostic Tests	1695	1300	394	879	1339	462	4213	754	816
Emergency Room Visits	69	29	40	28	46	6	41	12	35
DEA LAMBDA COEFFICIENTS				0.496	0.142	0.129	0.126	0.061	0.044
Based on DEA comparison of G234 with all 99 best practice specialists (G = generalist physician, S = specialist physician)									

Findings from multi-stage DEA analysis of generalist PCPs

Eight of the 47 generalists were found to be inefficient compared exclusively with BP specialists in stage 3. One of these is PCP G234, the generalist reported as inefficient in the initial analysis reported in Table 8-4. PCP G234 is now one of a group of generalists who were using more resources than the BP specialists, identified through this multi-stage DEA process. The results of this focused search for inefficient generalists with only specialists in their BP Reference Set are reported for PCP G234 in Table 8-7. The BP reference physicians for G234 now include only specialists: S340, S17, S78, S139, S127, and S245. The target resource level is lower than the previous analysis (Table 8-4) because the generalist physicians are no longer included to define the BP frontier. The potential resource savings are still substantial but are lower than the potential when generalists were included for all resources except for Therapy visits (see Table 8-8).

Table 8-8. Difference in Potential Savings for Generalist Physician

	Potential resource savings G234 compared with		Added potential savings - all physicians versus best practice specialists
	All physicians table 4 [a]	BP specialists table 7 [b]	[b-a]
Medical-Surgical Hospital Days	99	81	19
Ambulatory Surgery Units	3	2	1
PCP Office Visits	285	262	23
Referrals Resulting in a Specialist Visit	105	61	44
Mental Health Visits	38	22	16
Therapy Units	35	45	-10
Radiology, Lab and Diagnostic Tests	682	394	288
Emergency Room Visits	49	40	9
DEA efficiency rating	0.580	0.688	

Cost implications for managing specialist and generalist mix in the HMO

This finding suggests the possibility of a new cost control strategy that does not accept the belief that specialists are more costly in primary care. If the HMO wants to reduce costs by eliminating high resource use physicians, it should eliminate generalists like PCP G234 rather than the BP specialists. The common cost containment strategy – eliminate the BP specialists and retain the inefficient physicians like G234-- results in higher resource utilization. Ideally, the HMO should eliminate or retrain the inefficient generalists and specialists. HMOs may not have pursued this strategy because they cannot readily identify those specialists who are more efficient than the generalists. The multi-stage DEA approach suggested in this study may prove effective at identifying these high cost generalists and low cost specialists. Further study may be warranted before implementing this strategy to be sure the clinical implications and plan specific implications have been evaluated. In addition, the medical director might want review the case mix and quality differences among physicians to ensure that the data used in the analysis are appropriate.

Implications for HMO cost management

The key finding in this part of the study is that through use of multi-stage applications of DEA, we could identify generalists who used more resources than BP specialists. At least eight generalists used more clinical resources than specialists. In other applications, it may be necessary to rerun this analysis more times. Also there may not be such generalists in other HMOs. If we continued this multi-stage DEA application, however, we might uncover additional generalists who use more resources than specialists. This type of multi-stage DEA analysis may allow an HMO to identify generalists who use more resources than specialists. Further analysis of the mix of resources used would be needed to determine if the technically efficient specialist-PCP is also lower cost than the less efficient generalists. (If the DEA model applied incorporates relative cost data for the inputs, thereby addressing allocative and technical efficiency, it may be possible to identify generalists that are more costly than the BP specialists). Our findings indicate that HMOs might modify cost containment strategies focused on reducing specialists. Eliminating high cost, inefficient specialists *will* reduce costs. Beyond this, eliminating high cost, inefficient generalists may lower HMO costs more than eliminating low cost, BP specialists (as suggested in the example in Table 8-6). HMOs might benefit by developing more precise and sensitive methods of profiling high versus low cost physicians in each group. Such an analysis might be extended and refined to consider differences between each type of specialty. This may lead to other strategies to reduce costs.

8.3.5 Research Findings: Physician Group Practice Patterns - New Perspectives and Questions

The HMO we studied has physicians practicing in several provider groups. Alpha has separate contracts with each group. The reward system of sharing surpluses is based on group behavior rather than individual behavior. In this part of the study, we use the first stage DEA results to investigate whether the distribution of DEA scores within physician groups are similar. Since a physician's practice pattern could be influenced by his or her personal style, medical school training, and prior practice experience (Eisenberg, 1986), the groups might contain wide variations in practice style.

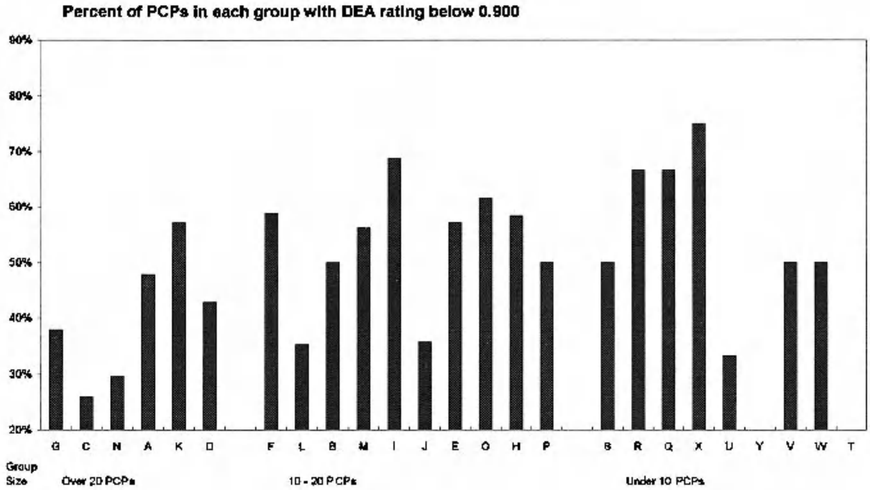
Potential differences in resource usage among groups

Table 8-9 lists the groups and the distribution of physicians in each group - BP, inefficient, and inefficient with DEA rating below 0.900. The Table 8. also reports the percentage of inefficient physicians in each group. The percentage of inefficient physicians differs between groups, suggesting that the group itself may influence inefficient practice styles. When these groups are sorted by size, the larger groups of 20 or more physicians range from about 41% inefficient to 71% inefficient. Similar swings are found in groups of 10 to 20 physicians. (The influence of the group and the patient population can be directly incorporated in the DEA analysis by adding categorical variables to the analysis as discussed in Rousseau and Semple (1993)).

Table 8-9. Distribution of Best Practice and Inefficient Physicians by Group

GROUP	TOTAL PCPs	BEST PRACTICE PCPs	INEFFICIENT PHYSICIANS DEA RATING BELOW 0.999	DEA RATING BELOW 0.900	% WITH DEA RATING BELOW 0.999		% WITH DEA RATING BELOW 0.900	
G	29	12	17	11	59%		38%	
C	27	16	11	7	41%	< low	26%	< low
N	27	15	12	8	44%		30%	
A	23	9	14	11	61%		48%	
K	21	6	15	12	71%	< high	57%	< high
D	21	11	10	9	48%		43%	
F	17	7	10	10	59%		59%	
L	17	9	8	6	47%		35%	< low
B	16	7	9	8	56%		50%	
M	16	5	11	9	69%		56%	
I	16	4	12	11	75%		69%	< high
J	14	8	6	5	43%	< low	36%	
E	14	5	9	8	64%		57%	
O	13	5	8	8	62%		62%	
H	12	2	10	7	83%	< high	58%	
P	10	4	6	5	60%		50%	
S	6	1	5	3	83%		50%	
R	6	2	4	4	67%		67%	
Q	6	2	4	4	67%		67%	
X	4	1	3	3	75%		75%	
U	3	2	1	1	33%		33%	
Y	3	3	0	0	0%		0%	
V	2	1	1	1	50%		50%	
W	2	0	2	1	100%		50%	
T	1	1	0	0	0%		0%	
TOTAL	326	138	188	152	58%		47%	

The bar graphs in Figure 8-9 further illustrate these variations. This reflects the percentage of inefficient PCPs and the percentage of inefficient PCPs with DEA ratings below 0.900 for each group. We presented this breakdown to Alpha's medical director, who had never considered this type of physician analysis. His curiosity about this analysis was almost as great as his interest in the original results that ranked physician efficiency. His interest was part professional and part in determining how his colleagues performed on this "physician report card".



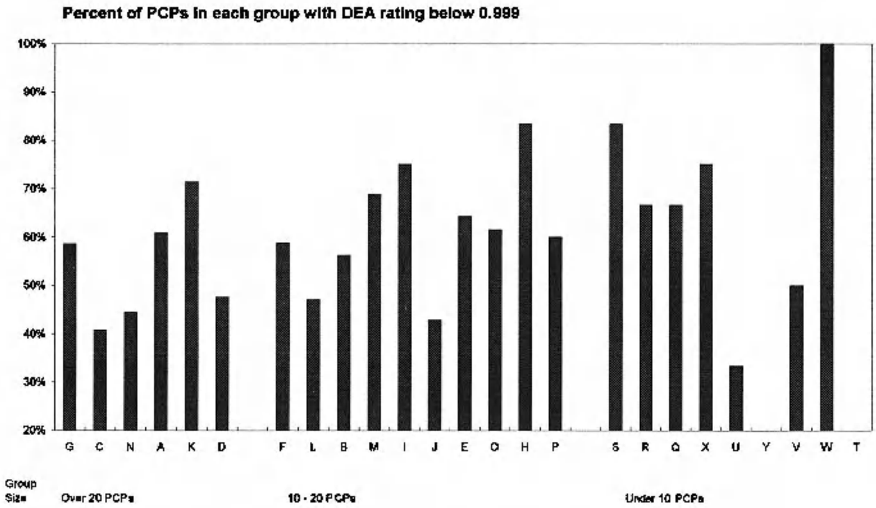


Figure 8-9. Efficiency Distributions for PCPs

Group differences raised new issues. Some groups with high concentrations of inefficient physicians were in low-income neighborhoods. This suggests that their patients – who have more complications due to their poor initial health – required more severe and costly care. These groups' higher resource utilization may be due to patient severity and not due to inefficient physician practice patterns. More sensitive severity measures would have helped ferret out the reason for greater resource utilization. Unfortunately, they were not available. Thus, these groups may be very efficient considering the patient mix. At the same time, each group has one or more BP physicians. These efficient physicians might serve as a more efficient care defining a practice pattern that other less efficient physicians in that group could adopt for in treating that group's patient population.

Alpha's medical director noted that one group had a different payment mechanism because it practiced exclusively in one of the hospitals used by the HMO. That group had the highest proportion of BP physicians. The director indicated that this group of physicians was not reimbursed for every resource used. Hence, the data on that group might understate actual resource utilization. The medical director's comments raised questions in our minds about the integrity of data collection across groups. Some BP physicians might be reclassified as inefficient if they were under reporting resources used. If these data were inaccurate, this could change the conclusions about which physicians were BP and inefficient. This potential data problem, first identified as a result of the DEA analysis, could cause problems in evaluating the physician costs regardless of the analytic technique used. Naturally, these data problems need to be resolved before the results of this analysis or any other analysis based on these data can be used as a basis for changing operating methods to reduce Alpha's costs.

The medical director indicated that the DEA results (Table 8-9) might suggest group practice styles, which influence Alpha's health care costs. In turn, the DEA results indicate the groups might reallocate the year-end "bonus" payments based on a physician's individual performance. Currently, the bonus distributed to groups is based on the overall plan surplus. Physicians within each group share that bonus in proportion to their base salary. No adjustment is made for individual contributions to the plan's profits. Given this, a BP physician in a weaker-high cost group is under compensated; an inefficient physician in a stronger-low cost group is currently over compensated.

Implications of group practice pattern variation on HMO costs and management

These findings suggest that HMOs might benefit from a better understanding of their group practice patterns and the reasons for differences in the efficiency of group practice patterns. This may allow them to identify the group patterns that should be adopted by the inefficient groups. HMOs should also consider that the inefficient groups might have some BP physicians. These doctors might serve as the model for the way patients can be treated at lower cost within the group's population, location, facilities, etc. Such physicians could train other doctors to practice more parsimoniously. In general, managing group behavior within HMOs may be a potentially fruitful new path to contain or reduce cost of care and DEA can help identify the physician practice patterns that can generate these cost savings.

8.4. CONCLUSIONS

In this physician-focused study we have introduced DEA as a tool to help evaluate physician utilization of primary care resources. Our analysis of 326 physicians' utilization behavior suggests the existence of a more "efficient" practice style. The conclusions may be summarized as follows:

1. A process to help inefficient physicians adopt BP patterns may lead to substantial resource savings in HMOs and other managed care organizations.
2. While specialists may have higher average costs than generalists, techniques like multi-stage DEA can locate specialists that are more efficient that high cost generalists. Instead of pursuing the current strategy of reducing specialists, HMOs should try to identify and retain the low cost BP specialists and generalists and try to reduce or retrain high cost specialists and generalists.
3. Different group practice patterns were identified which, when further analyzed, may suggest new avenues for managing costs and rewarding physicians in HMOs.

This study has several limitations that need to be considered in future applications of DEA to physician practice patterns. First, only crude measures of primary care case mix were used and no measures of outcomes or quality of care were available. A stronger test of quality of care would be to track the results of the treatments - a measure of the outcome as was done in Chilingerian (1989) and Chilingerian and Sherman (1990). The measure of treatments, visits, and other outputs is essentially a surrogate measure, which is used when outcomes are not measurable or too costly to capture. Although the study used age and gender to account for some utilization behavior, whether or not patients received an appropriate amount of care was not investigated. The possibility remains that much of the unexplained variation is the result of unmeasured case mix differences and/or quality of care differences. Interestingly, if Alpha's medical director is correct in assuming that specialists may care for sicker panels, our preliminary findings may be strengthened. To lend credibility to the frontier classification, a longitudinal study is needed to provide assurance that physicians practicing on the frontier always practice on the frontier.

Second, the definition of inputs and the degree of aggregation should also be re-evaluated. This would suggest whether a DEA analysis separating tests, treatments and other inputs into homogeneous sub-categories would alter the conclusions about individual physician productivity and the implications.

Third, and even more critical is the need to analyze an individual physician's overall patient satisfaction and outcomes in relation to treatment patterns and costs. This will help determine whether practice pattern differences can be identified at a more detailed clinical level. Specifically, results for individual BP and inefficient physicians need to be analyzed in detail by other physicians to verify and clarify the implications of these findings. Such interpretation would suggest whether excess use of resources was due to real inefficiency, or was justified by patient severity. In other words, are these PCPs really defining BPs in accordance with the philosophy of the HMO, or are such PCPs working in a more favorable operating environment, treating less complex cases, or holding back on needed resources? Just as there is no proof that "more is better," without better measures of primary care outcomes, efficient resource use may prove to be a poor practice pattern.

If these issues can be resolved, as has been successfully accomplished in other industries such as banking and government services, very specific prescriptions can be developed for HMO managers to reduce costs substantially. If the realized savings were only a fraction of the amount suggested in this study, the benefits to Alpha would be substantial. Individual HMOs that adopt this cost management strategy could become lower cost providers, more competitive and, directly or indirectly, more profitable. (Note that modest cost improvements in the managed care sector can have large absolute dollar impact, which suggests the value of further testing of this approach. If the managed care sector achieved one-half the potential cost savings identified in this one HMO, which is consistent with actual DEA results achieved other service organizations like banking, the U.S. health system cost savings could exceed $200 million per day).

Motivating physicians to adopt lower cost practice patterns

While the lessons from other service organizations are helpful, we recognize that the managed care industry has unique features and challenges. The ultimate question for the health care industry is: Even if we can clearly identify BP patterns of care, can physicians be motivated to adopt the patterns of care associated with least resource use? HMOs have not really exerted influence on physician practice patterns. Perhaps the approach of refusing to retain physicians who are below BP levels may succeed when and if HMOs become the best place to practice medicine. Alternatively, compensation systems more closely tied to "economic credentialing" - a reality in some managed care organizations - may provide the needed incentive. At this juncture, the question of how to motivate physicians to adopt low cost, BP patterns remains unanswered.

Research to test cost strategies and operationalize an approach to realize potential cost savings

We believe the most important element missing in this study is the absence of a review of the findings at the physician level by peers. Such a review would indicate whether the practice pattern differences are observable and due to excess resource utilization. If the study is to be used to change PCP practice patterns, the peer analysis would naturally make the DEA approach more credible and persuasive to the physicians being evaluated. Alpha did not have patient severity and health status data available in an easily accessible form, a necessary information component needed for this study and other purposes. It also was unable to commit the physician time that would have been required to sift through detailed patient records to gain these insights. An HMO with greater incentives to locate ways to reduce costs would be more likely to make this type of investment.

An alternative focus that was identified during this study was to analyze practice patterns for the illnesses that are most frequently treated and/or the illnesses that account for the highest total plan costs. This could be readily done with DEA and would be more easily analyzed from a clinical perspective to provide further insights into practice pattern differences.

Assuming we reach this level of confidence in clinical accuracy of the DEA results, the next step would be to test alternate mechanisms to induce inefficient PCPs to learn to use the BP patterns. This result might be accomplished through incentives, terminating contracts with high cost PCPs, and/or use of the economic credential in admitting PCPs to the plan. There is already evidence that more physicians are interested in joining HMOs than there is need for additional PCPs. The buyer's market for PCPs may allow the HMOs to select only those physicians that have high quality *and* low cost practice patterns

(Chilingerian (1990); Rosenthal (1994)). If further testing of this approach confirms the substantial cost containment benefits of the strategies identified in this paper, this area of research may constructively help answer the question posited earlier: How will health care providers deliver quality care with reduced revenues per patient and/or reduced total revenues?

This chapter is based upon "Benchmarking Physician Practice Patterns with DEA: A Multi-Stage Approach For Cost Containment" (Annals of Operations Research - Health Care Issue (1997) pp. 83 - 116), and –"Health Care Applications From Hospitals to Physicians, From Productive Efficiency to Quality Frontiers" (Chapter 17 in Handbook of Data Envelopment Analysis, Eds. W.W. Cooper, L. Seiford, J. Zhu – Kluwer 2004 pp. 481 - 538) both by Jon A. Chilingerian and H. David Sherman.[3]

APPENDIX A

Primary Care Inputs and Outputs Used in the DEA Analysis

Seven Outputs
Quantity of male and female enrollees 0-19
Quantity of female enrollees 20-39
Quantity of male enrollees 20-39
Quantity of female enrollees 40-59
Quantity of male enrollees 40-59
Quantity of female enrollees 60+
Quantity of male enrollees 60+

Eight Inputs
Quantity of Medical-Surgical Office Visits
Quantity of Ambulatory Surgery Procedures
Quantity of Medical-Surgical Hospital Days
Quantity of Mental Health Visits
Quantity of Emergency Room Visits
Quantity of Radiology, Lab and Diagnostic Tests
Quantity of Therapy visits
Quantity of Referrals Resulting in a Specialist Visit

[3] Jon Chilingerian is a Professor at Heller School for Advanced Studies in Social Welfare Policy, Box 9110, Brandeis University, Waltham, MA 02254-9110.

Chapter 9

GOVERNMENT PRODUCTIVITY MANAGEMENT

A Case Example

9.1. INTRODUCTION

Managers within the Government of Canada, like many governments, have demonstrated increased interest in methods to improve productivity and operations in general. This interest stemmed from: (i) pressures to assure that operations are well managed and (ii) concern that the techniques in use were not fully adequate or inappropriate for locating areas where improvements were possible.

An investigation into new, more sophisticated management approaches by the Deputy Minister of the Department of Supply and Services – Canada identified DEA as potentially suited to help improve and rationalize some elements of government operations. DEA had been demonstrated to be capable of locating ways to improve and manage productivity in several service environments including education, banks, health care, and computer service organizations.

Because of the limited experience with DEA in government applications, the Department of Supply and Services (DSS) [alternatively referred to as Supply and Services - Canada (SSC)] explores alternate pilot applications. Two key government organizations announced early interest in this pilot: The Staff of Regional Operations Section of SSC officers of the Management Practices Branch (MPG) of the office of the Comptroller General. For those more familiar with the US government structure, the DSS has similar responsibilities to the US General Accountability Office. The strong support of both senior appointees and senior civil servants provided a foundation that ensured the successful completion of this pilot.

The MPG developed a list of likely areas where the technique might he beneficial. This informal list included such diverse activities as the criminal justice system, taxation, health and welfare services, purchasing, and service operations like processing of government disbursements.

The pilot would focus on certain purchasing activities operated by SSC. This area of operation was selected for several reasons. First, it represented an area using significant resources. Also, the kinds of data necessary for DEA analysis were already collected in their management information system. Significant efforts had been expended to develop techniques to manage the purchasing operations, but the limitations of existing techniques suggested that other approaches could be beneficial as well. Finally, from an operational viewpoint, there were individuals on the SSC staff capable of learning and applying sophisticated management science techniques, and they could be immediately reassigned to manage this pilot study.

9.2. KEY ISSUES AND FINDINGS

The key issues and findings that resulted from applying DEA to this purchasing activity are described below. The remainder of this chapter focuses on the actual process of applying DEA in coordination with other techniques to manage productivity.

1. DEA has frequently been described as a complementary technique, one that must be used in coordination with other techniques to achieve productivity improvements. Application to purchasing services further confirmed this assertion and very explicitly explored the role of DEA vis-à-vis other techniques. 2. Introduction of a new method, and particularly a mathematically sophisticated technique like DEA, in an organization raises several management problems and issues. Examples of the issues that arose include: how to educate users, how to present the approach, how rapid adoption should proceed, and how does it relate to other techniques in use. The realization that DEA could be used for performance evaluation generated concern and emotional suspicions about the approach. Assertions that the technique was intended as a management aid rather than as a basis for penalties

did not mitigate this effect. As one manager succinctly stated, "If it can be used to measure performance, it will be used to measure performance." The mathematical complexity of this linear programming based method and its potential to evaluate performance resulted in considerable resistance. Another concern was that DEA might lead to reductions of department size and that such decisions might be based on more objective criteria than were used in the past. This might result in "changing the rules" and compromising political influence of some managers. These reactions are likely to occur to different degrees in all organizations. Grappling with these issues with the support of top management at DSS provided several insights about how to present, gain acceptance, and implement new sophisticated techniques.

3. Use of DEA and other techniques as management control tools raised questions beyond the capability of the methodology. The ideal frequency of use of DEA and other techniques and the gaming that will accompany their use were issues we confronted and these were fully addressed in the management literature.

4. A natural by-product of this project was the development of potential new ways to use DEA to increase its ability to discriminate efficient and inefficient service units. Specifically, weights were assigned to some of the inputs to reflect managements' understanding of some basic differences in the value of the inputs and outputs. This concept has since been widely developed and studied, referred to by researcher as weight constraints and cone ratios. The applications to DSS provided early evidence that this method of increasing the power of DEA generated insights that management believed to be meaningful and valuable.

5. Portions of the DEA results provided the impetus for changes in the commercial acquisition (CA) operations. Other insights from DEA that were acknowledged to suggest other likely areas of improvement were not used primarily due to political issues related to the branch offices and the managers involved.

9.3. PRODUCTIVITY MANAGEMENT FOR REGIONAL ACQUISITIONS AT DSS-CANADA

9.3.1 Scope, Objectives, and Performance Evaluation of the Commercial Acquisition Activity

The area selected to evaluate the productivity improvement possibilities of DEA was the group of regional offices for commercial acquisitions (CA). Total purchases of the SSC operation were about $7 billion per year, of which about $1.5 billion are the responsibility of the regional offices. Specialized groups within a Headquarters line organization handle particularly complex acquisitions. These specialized purchasing groups includes armaments and aerospace; marine; electronics and industrial systems; science and professional services; communication services; and, office automation. CA handles less complex but a more diverse range of purchases than these more specialized purchasing groups. For example, CA handles purchases of travel, food, fuel, clothing, medicine, laboratory equipment, armed forces uniforms, vehicles, and crash and fire trucks. Many of these purchases involve contracts that run for more than one year. As one senior manager noted, they buy "just about everything under the sun." Regional offices are responsible for only a part of these purchases. The CA responsibilities located in the department's headquarters (HQ) are not included in this study. The HQ handles larger, more complex CA contracts that are most effectively managed centrally, such as major fuel contracts and computer systems. HQCA has 435 staff members and handles 25,000 contracts per year. Regional offices have about 397 person years of staff and handle about 270,000 contracts per year.

A carefully delineated and very specific set of objectives had been established for CA purchasing in SSC against which performance was to be evaluated. The five objectives are as follows:

Sensitivity and Responsiveness to Customers: The primary purpose of CA is to serve as a purchasing agent to government. Its personnel were responsible for performing this role in a manner that satisfied the needs of the many government offices that rely on DSS for procurement on a timely basis at reasonable cost and quality. A key element used for evaluating this dimension is throughput time of the purchase contracts; the focus was on how well the CA met service delivery targets. Customer reports of satisfaction and dissatisfaction were solicited and monitored. In addition, managers would visit customers (government offices) to inquire directly about whether CA was meeting their needs.

Prudence and Probity: CA was expected to exercise reasonable judgment about the nature of contracts and purchase arrangements. Here, the issues related to the legality of various expenditures and whether government funds were being spent in a responsible and reasonable way. This dimension was monitored though ministerial inquiries, management audits and program evaluations completed on a regular basis.

Fairness and Equity to Suppliers: To meet customer needs, CA was expected to seek advantageous purchase arrangements. This objective was to be achieved with ethical business practices, which prevented taking advantage of or unfairly eliminating competition among vendors. Reviewing the extent to which competitive bids were solicited for purchases monitored this. This process also helps achieve price efficiency in government.

Economy and Efficiency: The time and resources used to contract for purchases was expected to be contained and/or minimized while meeting customer needs. Several ratios were used to monitor this activity, which were based on cost per contract, volume or contracts per person year, and the direct cost of buyers in relation to the indirect office costs such as those for support staff, supplies or rent.

National Objectives: A set of objectives and constraints exist to generate purchases of products made in Canada and promote small business development in Canada. Ratios used to monitor CA's compliance with these goals were percentage of small business suppliers to total purchases and percentage of Canadian content in goods purchased.

A review of the way management achieved these objectives indicates that satisfactory control was maintained with existing techniques in the following areas: responsiveness to customers, national objectives, prudence and probity and fairness to suppliers.

Two areas were determined to require additional management attention. Quality assurance was one; there was inadequate control of the extent to which efforts were made by CA purchasing agents to explicitly consider whether there should be and were competitive bids, whether small business suppliers were available and whether buying procedures and contract preparation met the established requirements and standards. Hence, a quality measurement and control program was initiated.

The area where there was least assurance that CA was well managed was on the economy and efficiency dimensions. Purchasing agents asserted professional judgment in contracting and each buyer handled a wide range of contract types. Because comprehensive standards did not exist, the purchasing operation exhibited many of the elusive problems associated with service businesses. The optimum efficient amount of resources needed was not known. The ratios used were not sensitive to the mix of activities in each office and consequently were interpreted only with respect to an individual field office; that is, an office would be compared with its performance in prior periods. Office managers met the concept of using these ratios to compare office performance with resistance because they serviced different volumes and mixes of contracts. Requests for added staff to meet service demands were based on overloads or increases in volumes without explicit analysis of purchasing transaction mix. Existing analytical ratios were not used for (and were not fully capable of) evaluating whether offices were efficient or whether they could and should improve their efficiency.

Data Envelopment Analysis was to be applied to determine if it could provide insights into ways to rationalize the evaluation of CA regional offices performance, improve efficiency, and reallocate resources among the various regional offices to improve productivity in CA activities.

Enhanced Weighted Ratios vs. DEA: One problem with the ratios used by CA for performance evaluation was that they were not sensitive to the volume and mix of purchase transactions. If the relative time needed for each different purchase type was available, a weighted output measure could be used and the ratios would reflect volume and mix. Advantages of weighted ratios would be:

1. Calculation would be simple.
2. Analysis would rely on existing data.
3. Monthly measures of performance would be available.

A set of such weights had been developed, but there were several concerns about proceeding in this manner:

1. The available weights were seven years old and there was some question of their reliability even when initially developed. These weights were established to set the transfer price to be charged government offices for CA buying services. A more complex purchase was charged out at a higher fee. The key objective was to charge out all CA costs to customers on some basis that differentiated the time and complexity of those services. The true complexity differences were not fully reflected by these weights and the only real requirement was that similar contracts be charged at similar fees. These fees could not readily be compared to a commercial or competitive fee by outside buying services.

2. Cost of developing meaningful weights was excessive based on prior experience, and it was not apparent that such an effort would yield meaningful weights. These costs included much personnel time and the process was viewed as highly disruptive.

3. Even with these weights, the identification of inefficient units would be subjective. The resulting ratios would reflect a combination mix, volume, and other characteristics that would be difficult to analyze and use to determine ways to improve productivity.

Project officers identified clear advantages of DEA over other methods like ratio analysis:

1. It obviates the need to develop standards.
2. It explicitly considers a complex and diverse set of activities.
3. It is conservative but fair and equitable in its approach so that field managers would ultimately accept it.
4. It is an inexpensive way to assign relative values as needed.
5. It could be applied using existing data (not really known at the outset of the project)
6. It is adaptable to an environment in which procedures change and which would otherwise require frequent recalculation of weights.

Their concerns about proceeding with DEA were:

1. Training the staff to interpret the results would be specialized due to the sophisticated math.

2. The way it locates inefficiencies is not highly visible, particularly for those with limited understanding of the technique.
3. Its complexity could reduce acceptability.

In addition, they were cognizant that all inefficiencies would not be located with DEA alone. However, in spite of its complexity and the limitations of existing methods, the potential benefits of DEA resulted in a commitment of top management to proceed. (A common element to many of the early DEA applications is there was a project manager who was intrigued and interested in exploring new potentially more sophisticated techniques to improve performance and relative comfort dealing with technical methodologies.)

9.3.2 Review of CA Activities in Identification of Outputs and Inputs

The project team reviewed the responsibilities, transactions, and other activities performed by regional offices with head office personnel and with managers in charge of several field offices[1]. At the same time, these managers were asked to indicate the specific types of services and transactions they were engaged in and the resources they used to provide these services. This stage served two distinct purposes. First, it helped to align the efficiency evaluation with the production process so that it would be as comprehensive as possible. Second, it clarified ways information from line managers regarding what they are responsible for and how different services are provided could be incorporated into the evaluation; this minimized the likelihood that the results would be discounted because important elements in their operations were overlooked or excluded.

Following this field review, relevant outputs and inputs were defined. The list of output transaction types and input resources that were to be used is presented in Table 9-1. The availability of data on these outputs and inputs was then investigated. The information system was found to capture data on most but not all of the outputs that were identified. Some compromises would be required, and a determination had to be made whether these compromises would result in an invalid or unreliable conclusion about the performance of these offices.

[1] The team was comprised of several SSC staff and David Sherman as an advisor on DEA.

Table 9-1. Outputs and Inputs Used to Evaluate Productivity of the Regional Offices

Outputs-Contract Services	Inputs
1. Normal Contracts - Multiple Source	1. Salary $'s
2. Normal Contracts - Single Source	2. Person Years
3. Request for Proposal/Quotation Multi Source	
4. Request for Proposal/Quotation Single Source	
5. Telephone buy - Multi Source	
6. Telephone buy - Single Source	
7. National Standing Offer	
8. Individual Standing Offer-Multi Source	
9. Individual Standing Offer-Single Source	
10. Draw Against Individual Standing Offer	

SSC staff decided to proceed with the outputs and inputs listed in Table 9-1. The outputs would include only purchase transactions and would reflect the number of each type of transaction completed over the prior fiscal year. The inputs would include only the personnel directly assigned to CA in each regional office and would be captured as one input measured in two ways, person years and dollar cost. DEA would be run twice using the personnel input first as person years and then as dollar cost.

9.3.3 Compromises in Data Used for CA

Outputs not captured in the information system were activities relating to training personnel to improve future performance, educating new personnel on basic procedures and improving the knowledge base of the personnel (trade meetings, etc). Excluding this type of data meant that the evaluation would assume that each office had the same proportion of these activities. If one office had high turnover and therefore high training costs, it would appear to be inefficient because its personnel would have produced fewer purchase transactions, even though they were producing more training. This apparent inefficiency might not be due to poor working procedures in processing contracts but rather due to time spent on training and added time spent by new staff in learning how to process these transactions. At the same time, the higher training cost may be due to other aspects of the way an office is managed which may be a source of lower productivity. The final productivity analysis would have to consider the circumstances of the inefficient offices and determine whether these compromises accounted for some or all of the identified inefficiencies.

Certain adjustments, revisions and other follow-up transactions, which are not identified as new transactions consume time but are not necessarily due to errors or mismanagement of the initial contract.

One option would have been to capture these as another set of transaction types, but this was not done and there were no data that reflected the frequency of this activity in each office.

Inputs included more than just personnel; however, personnel were the predominant input and were viewed by management as the proper focus. Hence, the assessment focused on personnel and excluded resources like supplies, telephone, and equipment. If one office had more automated systems than another, it might require fewer personnel. This issue would have to be reconsidered in the analysis phase to determine if there are cases where apparent inefficiencies are explained by other inputs and whether this has implications about how an office should be staffed and the types of equipment that should be used. Moreover, the productivity evaluation would focus only on use of personnel to produce services rather than use of all resources.

Personnel do not all have the same function, salary, and experience. By using just one input measure, person years (PY) or salary costs, the mix of personnel would not be explicitly considered. While these data were available. SSC staff chose to proceed with the single input measure for personnel. The option to redo the analysis was available and could be reevaluated at a later time. Salary cost was the first preference since that would relate most closely to the budget. PYs were to be used as well because it would provide some perspective about whether inefficiencies were due to elements that impact salary more than the number of personnel, such as geographic location and seniority. The question that would not be addressed is whether different personnel configurations, i.e. management, clerical, and buyers – would tend to result in more or less efficient regional office operation.

The SSC staff evaluated each of these compromises, and their judgment resulted in a decision to proceed. Understanding these limitations was critical to the success of this effort and helped ensure that the results were properly interpreted.

9.3.4 Time Frame

The initial analysis included data from twenty-eight offices over each of the past three years. Multiple-year data made it possible to ascertain whether there were trends in individual office or system-wide productivity and also whether offices appeared to be consistently efficient or inefficient.

9.3.5 Initial DEA Results

Table 9-2 reports the initial results of the CA office DEA review. These results were reviewed for apparent data errors, oversights, and other problems that would affect interpretation, first by project officers and then by regional office managers. All regional office managers also reviewed the results as a group at a conference.

Table 9-2. Initial DEA Efficiency Ratings (%) of Regional Offices

	Small Offices			Large Offices			
#	Yr. 1	Yr. 2	Yr. 3	#	Yr. 1	Yr. 2	Yr. 3
1	100	100	100	1	100	100	100
2	94.4	99.9	100	2	94.1	89.4	84.1
3	100	77.9	59.5	3	100	100	100
4	86.4	84.9	72.1	4	100	100	94.3
5	87.0	91.6	100	5	100	100	91.3
6	86.6	79.5	100	6	78.6	76.9	74.3
7	100	100	100	7	100	95.5	89.5
8	60.9	71.0	77.0	8	100	100	100
9	90.3	88.4	100	9	100	90.9	100
10	100	96.5	100	10	100	100	98.3
11	100	84.3	100	11	100	87.1	96.8
12	100	87.9	98.8	12	100	100	100
13	97.9	100	85.3	13	100	100	100
14	84.9	100	96.0				

Large and Small Offices Compared over a 3-Year Period. Input used is salary $'s

The initial results indicated consistent inefficiencies in several offices, no strong trends in system-wide productivity improvements or reductions, and some trends among a few offices. Several important insights arose from this process as described below.

Organizational Behavior Issues

Reaction to Complex Methodology: There was strong sentiment that DEA was too complex and unfamiliar at all levels of management. How does it work, is the evaluation fair, what will the results mean, and how will it impact my job and performance measure where questions raised by several managers.

Fairness of the evaluation due to its sensitivity to the output mix was a key factor that persuaded project officers of the potential value of this approach. In addition, it was seen by project officers to be more meaningful than simple ratio measures. Moreover, it was particularly well received as an evaluation tool that would counterbalance the widely used throughput time measure, which was viewed as a misleading measure of performance. Throughput time was accepted as one measure of quality of service in that it reflected the speed in which buyer requests were processed: it was not, however accepted as a measure of efficiency because it ignored mix, resources used, and complexity of the transaction.

By design, most regional managers were not trained in the underlying theory and mathematics to fully understand how the approach worked. They were, however, made aware of how the results would be interpreted. This experience indicates that DEA should not be rejected solely on the ground of its excessive complexity.

Presentations describing the *DEA approach* were designed to allow those at all management levels to question and probe the value of the technique. While the mathematical complexity was not eliminated, this process resolved most of the other concerns. Key points that arose related to the weights used to value transactions, the extent of DEA's ability to locate inefficiencies, and the way DEA addresses the mix of activities.

The weights assigned by DEA were understood to be calculated to make each office appear as efficient as possible when those weights were applied to all the 28 offices. Any other set of weights selected by managers would make that office appear as inefficient or more inefficient. This proved persuasive because it was perceived as "FAIR". In reality, this weighting could be viewed as too forgiving in that it tends to understate the true inefficiency present and gives the benefit of the doubt to each office.

DEA was described as incapable of locating all inefficiencies because it can only locate relative inefficiency. True inefficiency can only be measured when an absolute or optimum efficiency standard is available and such a standard was not available and could not readily be developed for CA. The question was whether the inefficiencies located would be sufficiently substantial and manageable to justify the cost of the process.

DEA approach to the input and output mix utilizes linear program formulations. Although well understood by a group of the SSC staff as well as some field office managers, it was viewed by others as a "black box" that magically generated an assessment of their efficiency. This black box effect was in itself disturbing to several managers who viewed it as an approach being imposed on them by headquarters. Exacerbating this concern were misunderstandings that arose when follow-up questions about DEA were explained by those not fully aware of DEA's capabilities, limits, and mechanics This resulted in some confusion and misguided expectations about DEA and added to the suspicions about and resistance to this approach.

Reaction to a New Technique: The DEA project was initiated by the staff at the head office and resulted in queries by field office management such as the following: Did the introduction of a new technique suggest that management was not already doing a good job at managing productivity? Or were they not finding adequate ways to evaluate productivity? Would the results require field office management to change their procedures? Would this reduce the ability of field managers to control their operations by introducing more centralized performance assessments? To some extent, line managers did not want a new approach imposed on them, and they were particularly concerned about an approach they might not fully understand and whose outcome they might not fully accept.

Field office managers assessed the validity, relevance, and weaknesses of DEA and they became convinced that the approach and its conclusions could not be rejected, disqualified, or discredited. They correctly concluded that it was not comprehensive and that other issues such as quality needed to be addressed as well. Quality was simultaneously reevaluated as will be described below.

Some managers had been investigating other ways of managing productivity of CA, and the DEA project accelerated their efforts to develop an alternative approach that would be comparable in capabilities to DEA, less complex, and more readily understood and managed. This alternative was basically an expanded ratio analysis and will be described at a later point. It had one advantage beyond simplicity – the ratio analysis allowed for continual monitoring on a monthly basis. In addition it proved to be complementary to DEA.

Evaluation of the Initial Data and DEA Application

Input-Output Data: The initial input-output specifications were considered to be adequate. No new insights into data problems arose at this stage.

Less Efficient Offices: DEA results pointing to the less efficient offices by senior management at SSC. These were found to confirm their perceptions about regional offices that were less efficient in a more quantitative and objective manner. No surprises were noted at this juncture; i.e., no offices that were noted with DEA as less efficient were strongly believed to be efficient.

Time Period: The primary objective of the analysis was to improve productivity and a reevaluation was proposed to further that aim. Based on three years of data, some offices were located as inefficient compared with other offices in past and subsequent years; but changes in the organization and its activities over the three-year period indicated that it would be more useful to focus on the question of which offices were inefficient compared with other offices in the most recent period and in a consistent organizational environment. Hence, only the most recent year's data were used in the revised analysis to restrict the analysis to units whose operating environment and management structure were consistent.

Office size: The regional managers raised some concern about the differences of offices across size dimensions, but this was not really a problem in the first analysis. No apparent scale economy bias was seen. For example, neither large offices nor small offices were consistently less efficient. Less efficient large offices were found to be inefficient compared with efficiency reference sets of small offices and vice versa. Nevertheless, to satisfy this intuitive concern, the revised analysis segregated large and small offices.

Contract Complexity: A key attraction of DEA was its ability to evaluate efficiency without assigning relative weights. Management noted, however, that while no reliable weights could be calculated, there were clear rankings of complexity. For example, a request for proposal purchase transaction requires more effort than a telephone buy. A modification of DEA analysis was proposed to reflect these rankings. This was done by adding a constraint that the weight assigned to telephone buys be less than that assigned to request for proposals. Several other constraints of this variety were added to incorporate known relationships into the analysis.

9.3.6 Refined DEA Results

DEA was rerun for the most recent year using the 10 contract types and the person-year and salary-dollar inputs, and with separate comparisons of large and small offices. In addition, the analysis was run without and with constraints reflecting knowledge of the relative ranking of resources needed for the contract. The transaction complexity rankings included were only those in which it was unarguable that substantially all the contracts would follow that rule.

The results of this refined analysis are presented in Tables 9-3, 9-4, 9-5, and 9-6. Additional interpretation of the results was also prepared as presented in Tables 9-7, 9-8 and 9-9.

Table 9-3. Final DEA Results: Large Offices Using Salary Dollars Input

#	1 Basic Model DEA rating No weight constraints	2 Enhanced Model DEA rating Weight constraints added*	Excess $ Salary Based on enhanced model**
1	100	100	0
2	100	80.3	$ 98,086
3	100	100	0
4	100	94.1	$ 60,062
5	100	95.1	$ 15,246
6	81.5	70	$ 261,390
7	91.2	84.3	$ 375,230
8	100	100	0
9	100	100	0
10	100	93.4	$ 23,338
11	100	100	0
12	100	100	0
13	100	100	0
14	100	100	0

* The discriminating power of DEA is increased by the inclusion of the output relationships; for example, the weights used to determine the efficiency level reflect the fact that the average level of effort required to do a normal contract will be greater than for a telephone buy or request for a proposal or quotation.
** Efficiency (%) x resources

Table 9-4. Final DEA Results: Large Offices Using Person Year (PY) Input

#	1 Basic Model DEA rating No weight constraints	2 Enhanced Model DEA rating Weight constraints added	Excess PYs Based on enhanced model
1	100	100	0
2	100	80.6	3.85
3	100	100	0
4	100	97.6	0.96
5	100	88.5	1.20
6	79	71.5	19.83
7	89.3	82.5	16.92
8	100	100	0
9	100	100	0
10	96.1	81.3	2.93
11	100	100	0
12	100	100	0
13	100	100	0
14	100	99.2	0.13

Table 9-5. Final DEA Results: Small Offices Using Salary Dollars Input

#	1 Basic Model DEA rating No weight constraints	2 Enhanced Model DEA rating Weight constraints added*	Excess $ Salary Based on enhanced model**
1	100	100	0
2	100	80.3	$98,086
3	62.4	100	0
4	100	94.1	$60,062
5	100	95.1	$15,246
6	81.5	70	$261,390
7	91.2	84.3	$375,230
8	100	100	0
9	100	100	0
10	100	93.4	$23,338
11	100	100	0
12	100	100	0
13	100	100	0
14	100	100	0

Table 9-6. DEA Final Results: Small Offices Using Person Year (PY) Input

#	1 Basic Model DEA rating No weight constraints	2 Enhanced Model DEA rating Weight constraints added*	Excess PYs Based on enhanced model**
1	100	100	0
2	100	100	0
3	71.1	66.1	0.678
4	72.3	66.8	3.373
5	100	100	0
6	100	100	0
7	100	100	0
8	75.9	70.5	2.705
9	100	100	0
10	100	100	0
11	100	100	0
12	100	100	0
13	100	100	0
14	97.6	96.2	0.079

Table 9-7. Summary of DEA Results for Inefficient Offices -- Potential Benefits of Eliminating Identified Inefficiencies

Inefficient Offices	Efficiency Rating ($ input)	Potential Annual Saving in $'s in Office Becomes Efficient	Potential Annual Person-Year Savings if Office Becomes Efficient
Large Offices#			
2	80.3%	$ 98,000	3.8
4	94.1%	60,000	0.9
5	95.1%	15,000	1.2
6	70.0%	261,000	9.8
7	84.3%	375,000	16.9
10	93.4%	23,000	2.9
Large office subtotal		$ 832,000	35.5
Small Offices			
3	60.2%	$21,000	0.6
4	73.0%	59,000	3.3
8	77.0%	43,000	2.7
13	97.0%	1,000	------
14	98.0%	1,000	0.1
Small office subtotal		$125,000	6.7
TOTAL POTENTIAL ANNUAL SAVINGS		$957,000	42.2

Tables 9-3 through 9-9 suggest the following:

Among the large offices, there are six that are relatively less efficient using the salary dollar input measure (Table 9-3): offices #2,4,5,6,7 and 10. Using the person year input, the same six offices and office #14 are found to be relatively less efficient (Table 9-4). The inconsistency with office #l4 was not investigated because the amount of inefficiency was small; i.e., there was an excess of 0.13 person years identified. While DEA tends to understate the real inefficiencies present, two of the large offices (#6 and 7) appear to be much less efficient than the more efficient offices.

Among the small offices, five were found to be less efficient: #3, 4, 8, 13 and 14 (Table 9-5), using salary dollars as the input measure. In Table 9-6, using PY as the input measure, only four of these were less efficient. Office #13 appeared as more efficient using PY as the input measure compare to the salary dollar input. Again, the source of this difference was not investigated because of the small degree of inefficiency located at this office.

Table 9-7 indicated that eliminating the inefficiencies identified could result in reduced cost of about $950,000/year or a reduction of 42 PYs. This is one of several possibilities.

Another possibility suggested by DEA is illustrated in Table 9-8 for office #7. Here, a reduction of resources used can be combined with an increase in transaction volume processed to make it as efficient as the more efficient offices. This range of alternatives to make office #7 more efficient is further described in Table 9-9.

Table 9-8. Illustration of Changes to Outputs Office 7 to Make it Relatively Efficient from DEA Results

Output-Contract Services	Actual Level of Activity	Composite of Relatively Efficient Offices	Increase in Outputs and Decreases in Inputs to make it as Efficient as Other Offices
1. Normal Contracts -- Multiple Source	1,971	1,971	----
2. Normal Contacts - Single Source	2,345	2,504	+159
3. Request for Proposal/ Quotation Multi Source	4,516	15,821	+11,305
4. Request for Proposal/ Quotation Single Source	6,357	6,357	----
5. Telephone buy - Multi Source	8,028	8,028	----
6. Telephone buy - Single Source	16,450	16,450	----
7. National Standing Offer	175	175	----
8. Individual Standing Offer - Multi Source	1,057	3,981	+2,924
9. Individual Standing Offer - Single Source	1,306	2,062	+756
10. Draw Against Individual Standing Offer	4,205	4,205	----
Inputs			
Salary $'s	$2,390,000	$2,180,000	-$210,000

As described in the chapter 2, a range of alternative paths is available to make the less efficient units in the group under study as efficient as the more efficient. These paths reflect mathematical relationships about the data and represent theoretical adjustments. At the same time, management may institute other more feasible organizational and operating changes that may not be apparent from the DEA analysis. The impact of such changes can be evaluated by using DEA as a sensitivity analysis tool to determine whether they are sufficient to make the unit relatively efficient.

Table 9-9. Demonstration of Analysis of DEA Results for Inefficient Offices

Large Office #7
Efficiency Rating = 84.3% (salary $ input)
= 82.5% (person year input)

Office #7 can become as efficient as the relatively efficient offices by:
A. Decreasing its inputs.
B. Increasing its output level,
C. Combination of input decreases and output increases.

A. Inputs reduced by 15.7% (which is 100-84.3%) would save $375,000 and make Office #7 relatively efficient.
B. Each output level increased by 18.6% [which is (100/84.34%)-100) with no reduction in input level would make Office #7 relatively efficient]
C. Combination of increased contract service activity and decrease in operating costs.

In this applied management setting, where there are political and organizational issues, where the underlying complexity of DEA is of concern to managers, and where DEA will not in itself specify exactly how efficiencies are to be achieved, a less aggressive interpretation of the results was adopted. The key insights derived for this application were which units were less and more efficient, what was the approximate magnitude of inefficiency, and which offices comprised the efficiency reference set for less efficient units. These data could have been used to designate the offices that would be focused on initially for productivity improvements and the offices that they would be compared with to determine how productivity could be improved.

The results were largely consistent with the initial analysis but were more focused. One view shows potential cost savings of over $950,000 per year or 42 PYs (see Table 9-7). This would represent a more than 10% reduction in staff in regional offices where there was excess capacity rather than an across-the-board cutback. Beyond this, the analysis would suggest that any request to increase staff would not be supportable by quantified data. As noted, this is but one interpretation; the actual savings that would be achieved could fall short of $I million, but it is also possible and more likely that the savings could exceed this amount because of the conservative way the DEA technique operates.

Several other approaches were used to understand implications of these results and locate ways to achieve productivity improvements. An informal review of a few offices was conducted to compare the procedures of relatively inefficient offices with those of efficient offices to assess the differences and their potential impact on productivity. This investigation was essentially an informal discussion of their operating procedures. Discussions with some regional managers and employees of these offices were conducted and observations of the workplace and the way buyer tasks and transactions were completed.

Two kinds of findings resulted from this effort. First, managers of less efficient offices identified differences in their offices that would make them less comparable to the best practiced offices. The differences may indeed cause an office to appear less efficient than it actually is. The question that needs to be addressed is whether this factor may have contributed to the amount of inefficiency noted. For example, do these differences really account for 5,10, or 15% excess resource usage and if they do is it really cost justified? The second set of findings was attributed to procedural and organizational differences that may impact morale and the process of completing a purchase transaction.

9.4. A NEW SYSTEM OF RATIO ANALYSIS TO CONTROL CA PRODUCTIVITY

The DEA productivity management project ran parallel to the development of a system aimed at tracking key operating relationships in each office. Its components included the following:

Efficiency measures

1. Personnel budget to actual by month.
2. Overhead or indirect operating costs as a percent of the direct personnel costs by office by month.
3. Overtime costs as a percent of direct personnel cost.
4. Salary costs as a percent of budget.

Effectiveness measures

1. Number of purchase transactions completed.
2. Number of transactions over x days in process (excessive throughput time.)
3. Error rate in contracting.

These analyses would eventually be used to compare an office against its performance in prior periods and against other offices. This type of ratio analysis provided conclusions about the offices that were highly inefficient similar to results obtained from DEA. The contract mix was not explicitly considered in this approach so the results would have to be adjusted qualitatively. This approach was also less specific about the location of inefficiencies and ways to improve productivity than DEA.

Rather than rely on a pure historical standard, this parallel program sought several managers' opinions to determine what their office performance levels ought to be. This was essentially a *best-practice review* that developed target performance levels based on a broad set of management experiences and operations.

Several capabilities not available with DEA were provided by this approach.

1. The analysis could be updated monthly to provide frequent indicators of problem areas and need for remedial action.
2. The ratios focused on certain key aspects that were not incorporated in the initial DEA analysis like overhead and overtime costs. (Such components could be included in a DEA analysis).
3. The types of data were relatively straightforward operating relationships that all managers could readily understand and analyze.
4. While DEA was unable to differentiate among the relatively efficient offices, ratio analysis provided significant insights about differences among several offices rated as efficient with DEA.

The key question is whether this parallel ratio based analytical approach, which supported the validity of the DEA results, would be effective in managing productivity. Best-practice reviews were used to develop standards based on qualitative and quantitative factors. Was that standard good enough or really efficient? Would substandard performance be explained by mix or other factors and would such performance be as prominently and objectively identified as occurs with DEA? To the extent that management initiative would be needed to determine which performance was substandard and requiring remedial action, would this system result in non-action as had occurred with other approaches? Finally, there remain questions about whether this approach is comprehensive enough to be used as the primary tool and even as the sole tool to monitor and manage performance: would productivity improvement possibilities be overlooked due to excessive reliance on this tool as might occur from excessive reliance on almost any single approach?

CA was informally committed to responding to the ratio analysis findings and to institute its analytic system to track progress from month to month. Tracking monthly progress can be more effectively done with their proposed use of operating ratios. DEA functions most effectively as a type of diagnostic X-ray, which locates problem areas, and the magnitude of the problem. For example, reapplication of DEA can pinpoint improvements as well as new problem areas where improvements are possible. In this pilot study, the application of DEA to the subsequent period indicated substantial productivity improvement in one office that was initially found to be the least efficient in the initial analysis.

9.5. SUMMARY OF RESULTS OF INITIAL PRODUCTIVITY REVIEW

Large office #7 was found to have an 84.3% efficiency level, which suggested potential savings of $375,000/year or 16.9 PY's if only inputs were reduced. Many other alternative adjustments were also ascertainable with DEA and specifically the direct path to the efficient frontier suggested saving of $210,000/year with the ability to have increases in certain transaction types. In as much as the input volume and mix were not manageable, the input adjustments were the primary focus.

DEA was one indicator of the area of greatest inefficiency. Ratio analyses (as described), which were the preferred approach of the field management, provided results that were consistent with DEA results and were therefore accepted as confirmation of the problem areas and of the accuracy of DEA.

9.6. SUBSEQUENT EVENTS AND THEIR IMPLICATIONS VIS-À-VIS THE DEA ANALYSIS

After completion of this project, several changes occurred. These events suggest the accuracy of the DEA results.

a. Managers were reassigned, resulting in moving the manager of one relatively efficient office to the large office *#7* (This was identified by DEA as the least efficient office).

b. Within six months of this time the DEA results were generated, the change in management in office *#7* gave rise to a decrease in personnel costs of about $530,000 per year or about 28 person years. These reductions were greater than the total anticipated with DEA for office #7, which is consistent with the nature of that technique; i.e., it will tend to underestimate real inefficiencies present. The decrease in PY's was done via attrition and by transferring PY's to the other activities of the same regional office. This suggests that the net decrease was somewhat smaller than noted above and that the other activities were either understaffed or are potential areas for future productivity improvement.

Several specific sources of the inefficiencies were identified by the new manager including (1) turnover due to the regional labor market, which attracts staff people away with higher wage rates, and (2) inconsistent processing procedures that were becoming more routinized in this process.

c. There was no change in the second less efficient office #6. Here there were no specific efforts to adjust operating productivity. Hence, while DEA suggested inefficiencies in office #7, which were eliminated in a short time frame, other inefficiencies suggested by DEA were not remedied in this same time period. DSS acknowledged the existence of real inefficiencies in office #7 but did not require any remedial changed in office #7 explicitly for political reasons relating to the rapport and influence the manager of that office had over activities in its geographic region. While more specific details were provided, they are not reportable and they don't change the net result of no action to improve office #7.

The DEA results were not sufficient to promote other improvements for several reasons.

1. The top down promotion of the DEA pilot was in itself a source of resistance to the approach.

2. The "black box" element in itself resulted in misinformation and misunderstandings about its purpose and capabilities, which presented another type of suspicion. For example, it was noted that several offices that were of very unequal productivity were all located as efficient and this was pointed to as evidence of erroneous conclusions from DEA. This, of course, is just an example of the inability of DEA (or any other available technique) to find all inefficiencies.

9.7. FIELD REVIEW FINDINGS

Field reviews indicated operating differences that could cause productivity differences among offices that might not be fully reflected in the data used for the DEA evaluation.

These include:

- unsolicited meetings by supplier
- unproductive phone calls
- file tracking and responding to client queries
- insufficient product specifications
- client relations
- supplier relations

Another set of differences, which might occur less randomly than the above and which may point to sources of inefficiency, were also identified:

- unexpected or temporary low business volume
- cyclicality of the business
- management style
- specialist vs. generalist approach
- buyer/support staff mixes
- synergism/common support vs. dedicated support staff
- overhead cost levels within regional offices

This second set of differences was specifically reconsidered in evaluating the DEA results. For example, was an inefficient office problem just one of an unexpected decline in volume? Separating large and small offices eliminated some of these fluctuations and using the full year tended to limit other seasonal issues, but these could have had an impact and still had to be considered.

Some differences in management style and employee attitudes were informally observed between two offices by one DSS staff person. While these differences could contribute to the differing levels of efficiency, these observations were not verified and consequently were not used to analyze ways to improve productivity.

Other differences noted between these offices, which may influence productivity, were the extent of availability and use of word processors, telex machines, computer terminal to locate suppliers that source specific materials, typing pool support and photocopying support.

Field review of this variety can be more focused and insightful using prior knowledge that particular offices were relatively less efficient or relatively more efficient. This is a key benefit resulting from analyses completed with techniques like DEA. With limited resources to evaluate field operations, management was able to deploy one project officer to selected offices to gain these insights.

9.8. QUALITY OF THE PURCHASE CONTRACTING PROCESS

The objective of the quality control program was to assure that at least a minimum acceptable quality level was achieved. Four dimensions of quality were identified: 1) contractual process used and documents produced; 2) sources (supplier) development errors; 3) supplier relations; and 3) customer relations. Since the contractual process used the majority of the CA resources, it was selected as the focus of the quality control program. A questionnaire about a contract was prepared which reflected the key attributes needed to meet established minimum contracting standards. Weights were assigned to each question to reflect the seriousness of shortfalls on each dimension. These weights were judgmental. Greater weight was assigned to deviations from legal requirements (Act of Parliament, Government regulation) and lower weight to poor purchasing practices. A "weighted error point average" resulted from the examination of a contract file and completion of the form. The presumption is that each contract and each office should score 100% because this was a minimum standard. Any ratings below 100% suggested need for improved quality. The assessment was to be done on a statistical sample basis. A sample of contract files of each office would be evaluated to get a measure of their quality of contracting.

An example of the questionnaire to evaluate quality is included in Table 9-10.

In the first three offices evaluated with this quality assessment form, two of the highly inefficient offices and one relatively efficient office were found to have quality levels below minimum standard. This provides some evidence that quality does not compensate for inefficiency and, indeed, low quality may result in inefficiency.

The quality assessment program has the hallmarks of quality management program for most services. Generally, there are several dimensions of quality. Management needs to select the key dimension to reduce the task to manageable size. This inevitably leads to potential criticism that certain dimensions of quality are absent or that surrogates of quality that are used don't reflect true quality. Such arguments need to be considered before finalizing the assessment program in interpreting the results.

The program adopted objectively located quality problems in the contracting dimension and explicitly did not incorporate other quality dimensions such as customer service. In the context of this project, the quality assessment was relatively less controversial. The quality measures identified offices where there was divergence from acceptable practices. It provides an assessment against standards accepted and understood by field managers. Moreover, the findings would tend not to have significant impact on their staffing and office operations except where irresponsible levels of quality were identified.

Table 9-10. Contracting Quality Assessment Checklist Example

A. REQUIREMENT	No. Of Occasions Commented as Unsatisfactory	Error Points
1. Were inadequacies or omissions in the requisition questioned?	1	2
2. Was there sufficient funding on contract award?	1	5
3. Were customer constraints not met or exceeded?	______	1
4. Were Security requirements followed?	______	5
5. Was the Requisition Acknowledged?	______	1
	Subtotal	_____
B. SOURCING (i) Sole or Single Source		
1. Should the requirement have been competed?	7	5
2. Was source justification weak or inadequate?	3	2
3. Was customer directed source challenged'?	2	2
4. Were efforts made to seek alternate sources?	3	1
	Subtotal	_____
(ii) Competitive		
1.Was correct method of requesting bids used (i.e. RFP, I/T. RFO, Speed Buy, Telephone/Telex)?	1	2
2. Was bid documentation package complete?	3	3
3. Were all sources invited to bid or explanation provided?	3	2
4. Were Suppliers rotated fairly?	______	2
5. Was sourcing adequate to obtain fair market price?	1	2
6. Was bid evaluation criteria satisfactory?	1	2
7. Were all bidders treated fairly and equally?	______	3
8. Was sufficient time allowed for bidding?	______	1
9. Were all bids open for acceptance at contract award?	______	3
10. If other than lowest bid accepted, is rationale justified?	______	3
	Subtotal	_____

Table 9-10 Contracting Quality Assessment Checklist Example (Continued)

C. PRICING	No. Of Occasions Commented as Unsatisfactory	Error Points
1. No or inadequate price support?	12	1
2. Correct price certification provided when required?	2	1
3. FST/Duty/PST position correctly stated?	2	1
4. Correct escalation clauses used as appropriate?	______	1
5. F.O.B. point(s) stated and priced?	2	1
	Subtotal	_____
D. AUTHORIZATION		
1. Contract Summary/Request on File7?	______	3
2. Contract correctly authorized?	1	3
	Subtotal	_____
E. CONTRACT DOCUMENT		
1. Is the contract in accordance with a) requisition (or is customer concurrence on file)? b) Supplier's offer? c) authority?	 3 3 ______	 3 3 3
2. Is contractor's name a correct legal entity?	2	1
3. Is contract correct type'? (YAR, YTA, YQA, PO)?	4	1
4. Has YAR type contract been acknowledged?	3	1
5. Is requirement, price, delivery and method of payment clear and complete?	6	3
6. Are all necessary terms and conditions in the contract clearly stated; e.g., sales taxes, duty, financial limitation, escalation, audit provisions, packaging and marking, shipping point, F.O.B. point, distribution of invoices, inspection authority, etc.	14	3
7. Was telex/telephone authority reflected in the contract?	______	1
8. Were contractual codings correct?	8	1
	Subtotal	_____
F. GENERAL		
1. Was file poorly documented?	6	1
2. Was Customer kept advised of changes?	______	1
3. Was telex of authority correctly worded?	______	1
4. Was ceiling price exceeded?	______	3
5 . Were late deliveries expedited?	1	2
6. Were payments made in accordance with contract?	1	4
7. Was there splitting of the requirement?	______	5
	Subtotal	_____
E. OTHER (Write in)		

9.9. ALTERNATIVE APPROACHES

Middle Management Training and Participation

The project described had one key requirement for success: top management support. At the same time, the communications to field managers resulted in resistance to the productivity improvement approach for reasons that were largely emotional. Some field managers believed this approach would be forced on them without considering the fairness, accuracy, and value of the conclusions. At least one field manager informally received misinformation about the nature and purpose of this approach, resulting in resistance to a misperception of this approach. While the intent was to pursue this approach as a pilot for careful scrutiny, the message was different.

Future use of complex methodologies like DEA requires more care in accurately introducing this to middle management and in the timing of that introduction. This analysis could have been completed without middle management involvement until the results were complete or they could have been involved more fully from the outset. Once they are introduced to this approach, they need to have clear explanations of the program and nature of the approach. This will not eliminate resistance but rather will force the debate to focus on real rather than imagined attributes of the approach. In this case, it could have been made clearer that it was in the field management's interest to help specify the inputs and outputs to fairly reflect their operations. It would also have been perceived as somewhat less threatening. More participation would also have made this seem less like an approach that would be adopted whether they like it or not. Moreover, the fairness of this technique, an attribute that is a key attraction to managers, would have been more evident. This would have balanced the concern about punitive effects with an appreciation for how this could improve operations.

One counter-example that provides some perspective on this is the use of a similar approach for high school (secondary school) principals in the U.S. While the principals who received data about their efficiency and effectiveness were not all fully aware of the technical underpinnings of this approach, the resistance has not proved to be a significant issue (Bessent et al. 1984). This is likely to result from two attributes. First, the training process about how to interpret the results and the reports themselves were designed to minimize ambiguity and misinformation. More importantly, the data is provided with the message that they should be used to help principals improve their productivity and that this is specifically not designed to penalize less efficient high schools. Indeed, a principal could choose to ignore this information. Peer pressure and pride are key motivators relied on to encourage use of these data. This more passive approach reduces potential resistance. It also does not assure that improved productivity will be achieved, which may be an unacceptable result considering the resources expended to provide these insights into productivity improvement possibilities.

Refinement of the outputs and inputs

The only input considered was personnel, with no differentiation between types of personnel. A more complete set of inputs could be used to provide a better understanding of where inefficiencies were the result of the mix as well as the volume of inputs. For example, the use of support staff vs. management vs. buyer to produce services was not considered. In addition, the use of purchased services, equipment and supplies could have been incorporated into the analysis to determine whether that had an impact on productivity.

Outputs used were found to be relatively comprehensive based on field discussions following the initial DEA review. Some outputs that could have been added are 1) training and other staff development activities and 2) adjustments to contracts which are not separately tracked but often are essentially the equivalent of new purchase transactions. This would serve to refine and more comprehensively measure the field office productivity.

PY vs. salary dollars

Results were essentially consistent using person years and salary dollars as the input. This suggests that the seniority and regional salary cost issue was not a key consideration.

Use of one input measure, however, did limit the value of the results, as it gave no hint about whether the management, clerical, or buyer staff was excessive for less efficient office. In addition, it provided only a general staff level for efficient handling of volume and mix rather than a profile of the efficient mix of personnel.

Integration of productivity management technique:

The emphasis on DEA at the central administration level and ratio analysis at the field office management level limits the coordination of these techniques to fully benefit from the attributes of each approach. Management needs to coordinate and reconcile the use and results of these techniques to identify changes and implement changes to improve productivity.

9.10. CONCLUSIONS

The combination of several approaches to assess productivity balanced with a program to consider quality and effectiveness resulted in specific areas and ways productivity could be improved. Some improvements were realized, suggesting that these insights were accurate. In this case, DEA was used as a lead instrument and was then coupled with ratio analysis, management reviews, best practice reviews and activity analysis to improve operations.[2]

This chapter includes materials from the monograph "Service Organization Productivity Management" by H. David Sherman (Society of Management Accountants of Canada, Hamilton, Ontario, 1988.) The Society of Management Accountants of Canada funded the study and publication. The monograph content was subject to extensive review by senior management of Supply and Services Canada and was accepted as an accurate representation of the project, process and impact of DEA. In addition, the senior manager of DSS independently issued a directive recommending expanded use of DEA to improve the efficiency of government operations.

[2] DSS formally declared that DEA was effective and that it should be adopted and used by more government departments. A note from the Deputy Minister of DSS asserted the value of DEA based on this pilot. At the same time, DSS did not continue to use DEA for a reason that is unlikely to be repeated in a non-government setting. On revisiting the department about 6 months after completion of the initial study, we found that that there was an unfilled position designated for an individual to pursue further DEA productivity analyses in the DSS. It was explained that this position was not filled at that time because the individual that would have been appointed based on civil service seniority was believed to be incompetent to complete such DEA analyses.

Chapter 10

MULTIDIMENSIONAL QUALITY-OF-LIFE MEASURE

An Application to Fortune's Best Cities

10.1. INTRODUCTION

A way to use DEA to evaluate quality is described in this chapter via an assessment of the relative quality of life of a group of cities using data reported in *Fortune* magazine. While this represents exploratory work, we include this because it represents an important development that can augment the benefits generated using DEA and it will enable readers to explore other ways to apply DEA to aspects of their organization. We encourage readers to explore the use of DEA to evaluate quality and to communicate your findings to other DEA researchers to allow them to continue to adapt DEA to manage productivity and quality.

Prior chapters emphasized our strong view that the power of DEA as a benchmarking technique to improve performance frequently requires that the impact on quality be explicitly considered. Chapter 7 offers several methods to include quality in the DEA applications. Quality is itself a multidimensional element. For example, hospital quality includes dimensions such as frequency of adverse events, delivery of care in accordance with hospital treatment guidelines, food quality, responsiveness of personnel, and training to administer care upon discharge from hospital. Quality is rarely a single dimension .

In addition to methods for incorporating quality in DEA, the topic of chapter 7, there is growing evidence that DEA can be a valuable tool to create a comprehensive index of quality. For example, multiple quality measures can be treated as multiple outputs in a DEA analysis. The input can be set as the number one. The resulting DEA scores will reflect the relative quality of the units being evaluated with the highest quality units having a score of unity and the lower quality units will have scores bellow one. Scores below one suggests the degree to which the units provide service below the highest quality units. DEA will also provide the reference set of units that point to the lower quality of service in units with ratings below one. In applications to quality, the reference set might be labeled quality reference set (QRS) in contrast to the efficient reference set (ERS). Use of DEA to create a quality index has been used by management in a 1500 bank branch application and in health care research applications. However, this work is in its early stages and no formal papers have been published evaluating the relative benefits of DEA over other methods of indexing quality. This chapter offers ideas to continue the development of DEA in evaluating and managing quality.

10.2. URBAN QUALITY OF LIFE ANALYSIS WITH DEA

"Quality-of-life" like "well being" or "social welfare", has a subjective or normative meaning (Slottje et al. (1991)). A typical measure of quality-of-life considers a finite set of measurable and objective attributes (factors) that can be weighted by some metric. It has been recognized that single dimension measures such as per capita GNP are too narrow to fully capture differences in the quality of life. Also, since these attributes have complicated and often indiscernable relationships with each other, multiple attributes are always necessary. Darton (1992) is the first to suggest using a multidimensional approach to the issue of economic well-being. The multiple attributes must be weighted in some objective and rational way in order to obtain an overall quality-of-life index. However, weighting the attributes of quality of life can be problematical, since we cannot know how individuals or groups weight the attributes of quality of life from which they derive utility. In the past, measures of the quality of life have been very narrowly constructed and have suffered from an assumption that the attributes are equally valued by assigning equal weights in practice. In some recent studies (e.g., Maasoumi and

Nickelsburg (1988)), among the feasible set of weighting schema, principal components and hedonic or instrumental variables estimating techniques have served as objective and rational ones. However, the weights obtained from these techniques sometimes are difficult to interpret.

A rich and very technical literature has also developed on using index numbers to measuring the quality of life (Diewert (1981)). Economists working on this problem have attempted (and continue to do so) to construct an ideal quality-of-life index, which satisfies a set of pre-determined axioms. Unfortunately, the result of this type of rigorous research has shown that such an ideal index does not exist and that the construction of one is not possible. This leads us to find alternative approaches to construct a quality-of-life measure. The multi-dimensionality of quality of life naturally raises the question of whether DEA can provide added insights. DEA has been applied to evaluate socio-economic performance (for example city and nation performance by Charnes, Cooper and Li (1989) and Golany and Thore (1997), and state of society evaluation by Hashimoto and Ishikawa (1993)). In particular, using DEA can circumvent the situation when information on how to weight multiple factors is not clear and even unknown.

We use 20 of Fortune's best cities in 1996 to illustrate how DEA can be used to measure the quality of life. In addition to Fortune's ranking, new DEA approaches are proposed to capture the difference in quality of life among the Fortune's best cities. For example, benchmarks are introduced into DEA model to (i) implicitly reflect factor tradeoff information and (ii) incorporate evaluation standards. The critical attributes related to the quality of life are identified for each city.

The next section provides a set of factors used by the Fortune magazine in selecting the best cites. DEA inputs and outputs are developed from these factors. We then develop some new DEA methods to measure the quality of life of 15 domestic and 5 international cities. The quality of life of the 5 international cities is compared to the best-practice of the 15 domestic cities. Conclusions are provided in the last section.

Table 10-1. Best Cities

		DEMOGRAPHICS			COST OF LIVING				BUSINESS	LEISURE			QUALITY OF LIFE	
Fortune's rank	City	1996 pop. in millions	Median household income	% of pop. with bachelor's degree	High-end housing price	Lower-end housing rental	Loaf of French bread	Martini	Class A office rental $/sq. ft.	Art museums	Public libraries	18-hole golf courses	Violent crime rate per 100,000	Doctors per 1,000
1	Seattle	2.2	$46,928	29.70%	$586,000	$581	$1.45	$4.50	$21	7	117	22	542.3	4.49
2	Denver	1.9	$42,879	29.10%	$475,000	$558	$0.97	$4.00	$14	5	60	71	595.6	2.79
3	Philadelphia	5	$43,576	22.70%	$201,000	$600	$1.50	$4.75	$21	25	216	166	693.6	3.64
4	Minneapolis	2.7	$45,673	27.00%	$299,000	$609	$1.49	$4.00	$24	6	131	125	496.5	2.67
5	Raleigh-Durham	1	$40,990	31.90%	$318,000	$613	$0.99	$4.50	$18	7	33	47	634.7	4.94
6	St. Louis	2.5	$39,079	20.60%	$265,000	$558	$0.89	$3.00	$18	10	104	62	263	3.4
7	Cincinnati	1.6	$38,455	19.90%	$467,000	$580	$1.25	$3.75	$20	4	71	94	551.5	2.8
8	Washington	4.6	$54,291	37.30%	$583,000	$625	$1.29	$3.75	$33	30	148	105	714.5	3.35
9	Pittsburgh	2.4	$34,534	18.80%	$347,000	$535	$0.99	$3.75	$17	8	124	112	382.1	3.66
10	DallasFort Worth	4.5	$41,984	27.10%	$296,000	$650	$1.50	$5.00	$18	3	98	77	825.4	1.96
11	Atlanta	3.5	$43,249	26.30%	$600,000	$740	$1.19	$6.75	$20	9	118	102	846.6	2.23
12	Baltimore	2.5	$43,291	23.30%	$575,000	$775	$0.99	$3.99	$18	8	102	45	1,296.30	4.02
13	Boston	3.2	$46,444	32.50%	$351,000	$888	$1.09	$4.25	$34	25	240	55	686.6	5.69
14	Milwaukee	1.5	$41,841	21.40%	$283,000	$727	$1.53	$3.50	$26	6	52	50	518.9	3.11
15	Nashville	1.1	$40,221	21.50%	$431,000	$695	$1.19	$4.00	$26	4	37	37	1,132.50	3.25
	International													
1	Toronto	4.3	21,000	17.80%	$290,000	$591	$1.00	$3.60	$31	90	178	122	415.4	1.91
2	London	6.9	14,535	11.0%	$700,000	$1,500	$0.95	$10.40	$83	300	443	95	430.6	3.65
3	Singapore	2.9	8,782	3.4%	$126,000	$2,128	$0.71	$8.52	$85	400	11	13	43.5	1.44
4	Paris	2.3	17,320	N.A.	$700,000	$1,200	$0.99	$8.87	$72	97	400	54	154.8	3.51
5	Hong Kong	6	8,000 (1)	7.1%	$769,000	$1,858	$1.20	$12.00	$119	15	61	4	156.3	1.26

10.3. QUALITY OF LIFE MEASURES

We use the data of 15 US domestic cities and 5 international cities as an illustrative example to demonstrate how DEA can be employed to measure the quality of life in a multidimensional construct and to provide additional information regarding the quality-of-life. Table 10-1 reports 13 factors (attributes) used by the *Fortune* magazine in selecting the best cities. These factors measure the cost of living, demographics, business and leisure. On the basis of these factors[1], we develop six DEA inputs and six DEA outputs as follows[2]. The DEA inputs represent *negative* evaluation items (smaller values are better and more desirable) and the DEA outputs represent *positive* evaluation items (greater values are preferred). The six inputs are High-end housing price (1,000 US$); Lower-end housing monthly rental (US$); Cost of a loaf of French bread (US$); Cost of Martini (US$); Class A office rental (US$/sq. ft.); and Number of violent crimes. The six outputs are Median household income (US$); Number of population with bachelor's degree (million); Number of doctors (thousand); Number of museums; Number of library; and Number of 18-hole golf courses.

10.4. MEASURING THE QUALITY OF LIFE ACROSS CITIES

This section demonstrates how the input-oriented CRS envelopment and multiplier models can be employed and modified to characterize the quality of life across cities. The discussion is carried out via four studies: (i) First, a method is proposed to integrate inefficiency represented by non-zero slack values into quality-of-life scores so that the quality-of-life gap with respect to the best-practice can be measured; (ii) Benchmarks (standards) are introduced into the input-oriented CRS multiplier model and are fixed as components of quality-of-life frontier for each domestic city under evaluation; (iii) A method is proposed to identify critical attributes with respect to the quality-of-life; and (iv) A DEA-based benchmarking model is used to

[1] See the November 1996 issue of Fortune for other factors including climate that are also used in selecting the best cities.

[2] For example, number of population with bachelor's degree is developed from the city population and the percentage of population with bachelor's degree. Note that some factor's units are changed.

measure the quality of life of international cities where the best-practice of domestic cities is given and used as a benchmark set. Each international city under evaluation can choose a proper subset of such benchmark set as the evaluation standard.

Quality-of-life rating and the identification of benchmarks

Note that the (input-oriented) DEA efficiency score (θ_o^*) measures the quality of life in terms of proportional reduction in all inputs of a city when outputs are fixed at their current level. i.e., θ_o^* is a radial (proportional) measure of quality of life. Slacks are likely present in the second stage analysis of DEA model.

We use the following index as a slack-adjusted quality-of-life measure

$$\theta_o^* - \frac{1}{m}\left(\sum_{i=1}^{m}\frac{s_i^{-*}}{x_{io}}\right)$$

The rationale is as follows. As pointed out by Cooper, Park and Pastor (1999), 1- θ_o^* provides a measure of "purely technical" inefficiency and $\frac{1}{m}\left(\sum_{i=1}^{m}\frac{s_i^{-*}}{x_{io}}\right)$ represents "average of input-mix" inefficiency. Thus,

$$1 - \theta_o^* + \frac{1}{m}\left(\sum_{i=1}^{m}\frac{s_i^{-*}}{x_{io}}\right) = 1 - \left[\theta_o^* - \frac{1}{m}\left(\sum_{i=1}^{m}\frac{s_i^{-*}}{x_{io}}\right)\right]$$

provides the total input inefficiency for a city under evaluation. The total inefficiency in the above equation actually measures the (input-oriented) quality-of-life gap between the best-practice and the city under evaluation. If the slacks are not adjusted, then the evaluation of quality of life is not being compared to the best-practice unit.

Table 10-2 reports θ_o^* and its slack-adjusted value. Under the assumption of CRS, four US cities (Cincinnati, Atlanta, Milwaukee and Nashville) are not on the quality-of-life frontier. Note that Milwaukee has a better quality-of-life rating than Cincinnati in terms of θ_o^*,whereas the adjusted θ_o^* shows an opposite result. This further indicates the necessity to integrate slacks in quality-of-life rating in order to fully and correctly characterize the quality of life.

From the benchmarking point of view, it is also important to identify the benchmarks for non-frontier cities. We can obtain this type of information via the non-zero optimal lambda values in the input-oriented CRS envelopment model. Table 10-3 shows the benchmarks for the four non-frontier cities under CRS. For example, the benchmarks for Cincinnati are St. Louis, Minneapolis, and Pittsburgh[3].

Table 10-2. DEA Results for 15 US Cities

Fortune's rank	City	CRS score	slack adjusted
1	Seattle	1	1
2	Denver	1	1
3	Philadelphia	1	1
4	Minneapolis	1	1
5	Raleigh-Durham	1	1
6	StLouis	1	1
7	Cincinnati	0.97697	0.88512
8	Washington	1	1
9	Pittsburgh	1	1
10	Dallas-Fort Worth	1	1
11	Atlanta	0.95717	0.77401
12	Baltimore	1	1
13	Boston	1	1
14	Milwaukee	0.99406	0.85128
15	Nashville	0.78979	0.77305

Table 10-3. Benchmarks for non-Quality-of-Life Frontier Cities

Non-frontier city	Benchmarks*							
Cincinnati	St. Louis	0.44555	Minneapolis	0.08091	Pittsburgh	0.50233		
Atlanta	Washington	0.19715	Philadelphia	0.28107	Pittsburgh	0.01883	Denver	0.4582
Milwaukee	Philadelphia	0.02551	St. Louis	1.04224				
Nashville	Denver	0.09926	Seattle	0.04965	Washington	0.10383	St. Louis	0.7165

* The number next to the benchmark city represents the optimal lambda value in the input-oriented CRS envelopment model.

Incorporation of benchmarks and value judgment

The above analyses reflect the natural structure of the data set. However, as we can see, about 75% (four of the 15) of the US domestic cities represent the best-practice of quality of life. It has been recognized that this is caused by having too many inputs and outputs and the weight flexibility in DEA models. Numerous methods have been proposed to reduce the number of frontier DMUs if this is perceived as necessary. For example, we may (i) incorporate some

[3] In DEA, the performance of these three benchmark cities dominates that of Cincinnati.

weight restrictions, as in ARs discussed in chapter 5 or (ii) use the preference structure model of Zhu (1996). Note that all these methods need additional explicit information on tradeoffs among inputs and outputs. Unfortunately, the current study does not have access to this type of information.

One might argue that one could reduce the number of DEA inputs/outputs by using some statistic techniques, e.g., correlation analysis. But this may affect the comprehensiveness and accuracy of the measure in a sense that not all dimensions are considered even though some dimensions may be strongly correlated. Therefore, the current study seeks an alternative way to implicitly express the tradeoff information and further to reduce the number of frontier DMUs.

The Fortune magazine ranks Seattle, Denver and Philadelphia as the top three best cities to balance work and family life. Note that for a city under evaluation, the input-oriented CRS multiplier model determines a set of referent frontier cities, which are represented, by a set of binding constraints. These frontier cities actually form the benchmark-set for a particular city under evaluation. The tradeoff information is represented by the efficient facets constructed from these frontier cities. Therefore, setting (fixing) the top-three Fortune cities as benchmarks in our DEA analysis can implicitly expresses tradeoff information. We proceed as follows.

Let set $\boldsymbol{B} = \{ DMU_j : j \in \mathbf{I}_B \}$ be the benchmark set. In this case, $\boldsymbol{B}$ = {Seattle, Denver, Philadelphia}. We then modify the input-oriented CRS multiplier model to the following linear programming problem[4]

$$
\begin{aligned}
& \max \sum_{r=1}^{s} u_r y_{ro} \\
& s.t. \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} = 0 \qquad j \in \mathbf{I}_B; \\
& \qquad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} \le 0 \qquad j \notin \mathbf{I}_B; \\
& \qquad \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \qquad\qquad u_r, v_i \ge 0.
\end{aligned}
\tag{10.1}
$$

[4] This is called fixed-benchmark DEA model in Zhu (2003) and Cook, Seiford and Zhu (2004). Zhu (2003) offers an Excel Add-In software for this type of models.

By applying equalities in the constraints associated with benchmark DMUs (cities), model (10.1) measures a city's quality of life against a reference set containing set ***B***. i.e., the top-three Fortune cities must be used in constructing the efficient facets. The equality-constraint associated with set ***B*** implicitly represents tradeoffs among various inputs/outputs.

Note that model (10.1) may be infeasible because (i) the cities in set ***B*** cannot be fit into a same facet when the number of these cities is greater than m+s-1, where m is the number of inputs and s is the number of outputs, and (ii) the cities in set ***B*** construct an inefficient (dominated) facet. Case (i) can be avoided by selecting benchmark cities such that the number of selected cities is less than m+s-1. Case (ii) can be circumvented by modifying model (10.1) into

$$
\begin{aligned}
z_o^* = \max & \sum_{r=1}^{s} u_r y_{ro} \\
s.t. \quad & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} = 0 \qquad j \in \mathbf{I}_B; \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} \le 0 \qquad j \notin \mathbf{I}_B, j \neq o; \\
& \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \quad u_r, v_i \ge 0.
\end{aligned}
\tag{10.2}
$$

If a city "o" dominates cities in set ***B*** (case (ii) and model (10.1) will be infeasible), we have $z_o^* > 1$, where z_o^* is the optimal value to model (10.2).

Table 10-4 reports the results from model (10.1). Seven cities (Minneapolis, St. Louis, Washington and Pittsburgh and the cities in set ***B***, Seattle, Denver, and Philadelphia) are now on the set ***B*** adjusted quality-of-life frontier, indicating that these seven cities are best-practice cities in terms of quality-of-life. Four frontier cities (Raleigh-Durham, Dallas-Fort Worth, Baltimore and Boston) are no longer on the set ***B*** adjusted frontier. The scores for the original four non-CRS-frontier cities are also dropped. We note that, based upon the scores obtained from model (10.1), again Cincinnati has a better quality-of-life status than Milwaukee.

We may divide the 15 cities into two groups on the basis of model (10.1)'s ranking (Table 10-4): the top-nine Fortune cities and the remaining six cities. In this case, after incorporating the benchmarks, our DEA method provides a very consistent ranking with Fortune's, although DEA ranking has ties in the top-nine Fortune's best cities. This further justifies the DEA results with the Fortune's ranking, and vice versa.

Table 10-4. Benchmark-adjusted Result

Fortune's Rank	City	Scores	Rank
1	Seattle	1*	1
2	Denver	1	1
3	Philadelphia	1	1
4	Minneapolis	1	1
5	Raleigh-Durham	0.93893	8
6	StLouis	1	1
7	Cincinnati	0.89768	9
8	Washington	1	1
9	Pittsburgh	1	1
10	Dallas-Fort Worth	0.87879	10
11	Atlanta	0.81361	13
12	Baltimore	0.84575	12
13	Boston	0.86822	11
14	Milwaukee	0.80851	14
15	Nashville	0.75470	15

* The score for a benchmark is obviously one.

Critical quality-of-life factors

Note that four cities (Cincinnati, Atlanta, Milwaukee and Nashville) are not on the best quality-of-life frontier, therefore every factor is very important in order to reach the frontier. However, we can use the following DEA model to determine the minimum change required on each individual input in order to reach the frontier[5].

$$\begin{aligned} &\min\ \theta_o^k \\ s.t.\ &\sum_{j=1}^{n} \lambda_j x_{kj} = \theta_o^k x_{ko} && i = k \\ &\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} && i \neq k \\ &\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} && r = 1,2,\ldots, s; \\ &\lambda_j \geq 0 && j = 1, \ldots, n. \end{aligned} \tag{10.3}$$

[5] A similar model can be obtained to determine the minimum change required for each individual output in order to reach the frontier.

Let $\theta_o^{i*} = \max_k \{ \theta_o^{k*} \}$, then this particular ith input gives the shortest path to the quality-of-life frontier for city "o". We say that this input is critical for the city's quality-of-life. Table 10-5 reports the critical quality-of-life factors for the four non-frontier cities. The household income appears to be the critical output for Cincinnati. The housing rental is the critical input for Cincinnati and the housing price is the critical input for Milwaukee.

Table 10-5. Critical Quality-of-Life Factors

Fortune's Rank	City	Critical Factors*
1	Seattle	{housing rental}, {household income}
2	Denver	{housing price, housing rental, office rental}, {household income}
3	Philadelphia	{cost of living, business}
4	Minneapolis	{cost of living, business}
5	Raleigh-Durham	{household income}, {household income, bachelor's degree, doctors}
6	StLouis	{French bread, martini}
7	Cincinnati	{housing rental}, {household income}
8	Washington	{cost of living, business}
9	Pittsburgh	{French bread, martini}, {violent crime}, {leisure}
10	Dallas-Fort Worth	{housing price, housing rental, office rental}, {office rental}, {bachelor's degree}, { bachelor's degree, doctors}
11	Atlanta	{French bread}, {cost of living, business}
12	Baltimore	{housing price, housing rental, office rental}, {office rental}, {French bread, martini}, {household income}
13	Boston	{cost of living, business}
14	Milwaukee	{housing price}, {cost of living, business}
15	Nashville	{French bread}, {cost of living, business}

*The critical factors are identified in a different way for the four non-quality-of-life frontier.

For efficient cities, critical factors are identified with respect to the following rule: the quality-of-life classification of a city changes if the magnitude of (some) quality-of-life factor(s) changes. Accordingly, model (10.3) is modified as

$$\begin{aligned}
&\min \theta_o^I \\
s.t.\ &\sum_{\substack{j=1\\ j\neq o}}^{n} \lambda_j x_{ij} = \theta_o^I x_{io} && i \in I \\
&\sum_{\substack{j=1\\ j\neq o}}^{n} \lambda_j x_{ij} + s_i^- = x_{io} && i \notin I \\
&\sum_{\substack{j=1\\ j\neq o}}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} && r = 1,2,\ldots,s; \\
&\lambda_j \geq 0 && j \neq o.
\end{aligned} \tag{10.4}$$

where I represents an input subset of interest. A similar model can be obtained based upon outputs. Model (10.4) may be infeasible for some I and further the infeasibility means that input changes associated with set I across all cities do not change the efficiency status of city "o". Consequently, we here can use the infeasibility information to identify the critical quality-of-life factors. That is, if model (10.4) is feasible, then factors in set I are important to the quality-of-life of a specific efficient city[6].

Model (10.4) is applied when set I only has one of the six inputs. Also, some combinations of inputs in set I are considered. For example, set I = {cost of living} = {housing price, housing rental, French bread, Martini}, set I = {cost of living, business} = {cost of living, office rental} and set I = {leisure} = {museums, libraries, golf courses}, etc.

Table 10-5 also reports the critical factors for the best-practice cities. For example, housing rental and household income are two critical factors to Seattle's quality-of-life. Note that the median household income for Seattle is almost \$47, 000, ranking only below Washington, D.C. and San Francisco. If this output is decreased (along with increases in other cities), Seattle may not be on the quality-of-life frontier anymore, not to mention the number one rank.

Cost of living and business office rental together constitute a critical factor for Boston's quality-of-life. Note that Boston's cost-of-living index is very highly ranked among the nation's largest cities.

Finally, we shall point out that the critical factors here are identified under a relative basis, since DEA compares relative performance

[6] See Chen and Zhu (2003) for detailed discussion on how to use DEA to identifying critical measures.

within a group of cities. If some existing cities are excluded or new cities are added, then the corresponding results may vary.

Measuring the quality-of-life of international cities

Since we have the quality-of-life frontier of the 15 domestic cities, we may compare the quality-of-life of each international city to this existing frontier. Troutt, Rai and Zhang (1991) and Seiford and Zhu (1998b) develop a DEA-based method to determine whether a new case is acceptable compared to the existing standard. As an extension to Seiford and Zhu (1998), we use the following linear programming to measure the quality of life for each international city

$$
\begin{aligned}
&\min \theta_{new} \\
s.t.\ &\sum_{j\in E}^{n} \lambda_j x_{ij} + s_i^- = \theta_{new} x_i^{new} && i = 1,2,\dots,m; \\
&\sum_{j\in E}^{n} \lambda_j y_{rj} - s_r^+ = y_r^{new} && r = 1,2,\dots,s; \qquad (10.5)\\
&\lambda_j \geq 0 && j \in E.
\end{aligned}
$$

where x_i^{new} and y_r^{new} are the ith input and rth output for a new city – an international city, respectively.

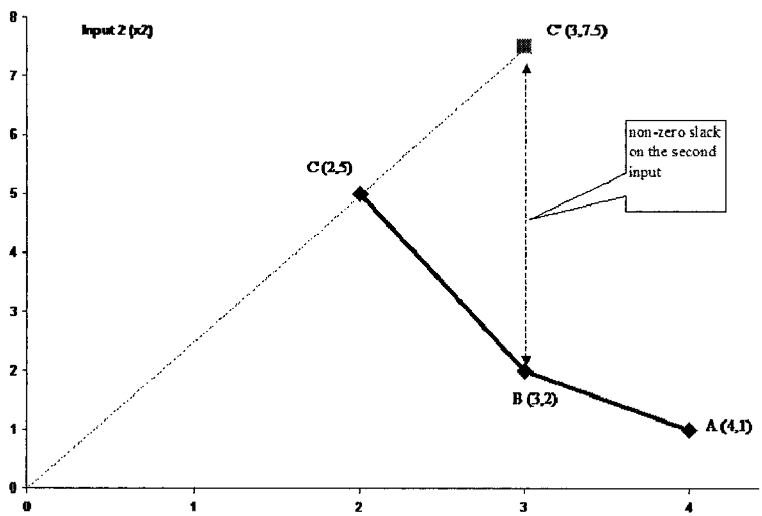


Figure 10-1. Super-efficiency and Non-zero Slacks

On the basis of Cook, Seiford and Zhu (2004), when assuming all data are positive, we have (case 1) $\theta^*_{new} \geq 1$ indicating the performance of this new city is at least as good as that of frontier cities represented by set E, and (case 2) $\theta^*_{new} < 1$ indicating the performance of this new city is worse than that of frontier cities represented by set E.

Now, consider three DMUs, A, B and C, with two inputs and a single output pictured in Figure 10-1. If we set E = {A, B} and solve (5) for C, we obtain

$$\theta^*_C = 1.5,\ \lambda^*_B = 1,\ s_2^{C-*} = 5.5 \text{ and all other variables zero}$$

Note that $s_2^{C-*} = 5.5 > x_{2C} = 5$. This is due to the fact that C is projected onto C′ which is on the extended DEA frontier of the two remaining DMUs A and B. (DEA frontier is the line segment AB.) However, C′ should be rated as inefficient (among A, B and C′) because of the non-zero slack value in its second input. In order to consider the inefficiency represented by non-zero slack values, we establish the following linear programming problem after solving model (10.5):

$$\begin{aligned} &\min \frac{\sum_{i=1}^{m} \frac{s_i^{o-}}{x_{io}}}{m} \\ s.t.\ &\sum_{j \in E} \lambda_j x_{ij} + s_i^{o-} = \theta^*_{new} x_{io} \qquad i = 1,2,\ldots,m; \\ &\sum_{j \in E} \lambda_j y_{rj} - s_r^{o+} = y_{ro} \qquad r = 1,2,\ldots,s; \\ &\lambda_j \geq 0 \qquad j \in E. \end{aligned} \tag{10.6}$$

where θ^{s*}_{new} is the optimal value to (10.5) and is fixed in (10.6). Model (10.6) maximizes the average input slack mix. We now define

$$\Theta_{new} = \theta^*_{new} - \frac{1}{m}\left(\sum_{i=1}^{m} \frac{s_i^{o-*}}{x_{io}}\right) \tag{10.7}$$

Θ_{new} can be used to represent the average performance gap between a unit on the best-practice frontier and a city under evaluation. For point C in Figure 2, we now have

$$\Theta_C = \theta^*_C - \frac{\frac{s_1^{C-*}}{x_{1C}} + \frac{s_2^{C-*}}{x_{2C}}}{2} = 1.5 - \frac{0 + \frac{5.5}{5}}{2} = 1.05$$

The rationale for using (10.7) is that if the slacks are not considered, the benchmark score over-states the performance of a city under evaluation. We illustrate this by the following application.

Table 10-6 reports the results from models (10.5) and (10.6). On the basis of model (10.5), Hong Kong is the only under-performing city compared to the best-practice of 15 US cities. However, if we use the slack-adjusted scores in (10.7), Singapore becomes another under-performing city in terms of the US best-practice of quality of life. Table 10-6 also provides the average input values of the 15 US cities. Note that Singapore has a very small value on violent crime compared to the average of US cities[7]. By the proportional change on all inputs in model (10.5), this small value leads to a large benchmarking score (optimal value to model (10.5)) with a zero slack on violent crime. As a result, the mechanism of model (10.6) yields large slack values on the other inputs for Singapore. As a matter of fact, all non-zero slack values are much greater than the original input values. Therefore, the original benchmarking score is biased by focusing on the violent crime only. The slack-adjusted benchmarking score given by (10.7) on the other hand balances all the inputs and yields an improved result.

10.5. CONCLUSIONS

In measuring the quality of life of cities, the November 1996 issue of Fortune "threw all the information into a database, contacted 650 high-ranking executives, and came up with a list of 20 winning cities". Obviously, it is a tedious task to develop measures to balance numerous factors that contribute to the quality of life. Not only is it because of the multi-dimensionality, but also the often-unknown relationship among various quality-of-life factors. This chapter illustrates that by using DEA, one is able to develop a multi-dimensional quality-of-life measure without *a priori* knowledge on the factor-relationship.

Some new DEA developments are used to capture the multi-dimensionality of quality of life and to measure the practical comfort for living. The study also offers a way to identify critical quality-of-life factors for each city. Such new information will be important to maintaining the best quality-of-life status.

The city quality of life analysis suggests new ways to incorporate benchmarks into DEA models in two ways. One is to use all the

[7] The average violent crime of the 15 US cities is almost 14 times of that of Singapore.

benchmarks in constructing the efficient facets for each city under evaluation. This allows the implicit incorporation of tradeoff information on various quality-of-life factors. The other is to allow each city under evaluation to select a subset of benchmarks. This new method is particularly suitable for measuring the quality-of-life of international cities, since each international city is measured against the same best-practice frontier (standard). Plus, since the number of outputs and inputs in the current study is much larger than the number of international cities, only one city (Hong Kong) was inefficient if the standard DEA model is used. Thus, the standard DEA model is unable to discriminate the quality-of-life among international cities. The benchmarking DEA model offers an alternative way to increase the analytic power of DEA with the introduction of external validated constraints.

Fortune ranks the 20 best cities. Our objective was to offer an alternative perspective on how to measure quality-of-life. Our hope is that readers will consider the way quality was evaluated in this study of cities and develop new ways to adapt this to the quality evaluations in your organization.

This chapter is based upon Zhu, Joe, A multidimensional quality-of-life measure with an application to Fortune's best cities, Socio-Economic Planning Sciences, Vol. 35, Issue 4 (2001), 263-284, with permission from Elsevier.

Chapter 11

HEDGE FUND PERFORMANCE EVALUATION

11.1. INTRODUCTION

Hedge funds are alternative investment vehicles that aim to provide absolute returns in any market environment with low positive or even negative correlation to traditional stock and bond markets. Hedge funds have gained popularity and acceptance by institutional investors for diversifying traditional stock and bond portfolios. While these funds have been primarily available to wealthy sophisticated investors, they are increasingly available to a wider set of investors and their operations are being more closely scrutinized by securities regulating authorities such as the U.S. Securities and Exchange Commission. Mutual funds that focus on direct investments in stocks and bonds are widely evaluated and there are many accepted methods of measuring their total return to shareholders. Hedge fund investments include stocks and bonds but also include other financial instruments that are by their nature focused on balancing risk and return. Consequently, the traditional total return measure used for mutual funds is not sufficient for hedge funds and other techniques are needed to provide a fair measure of comparative performance. This chapter suggests ways DEA can provide insights into hedge fund performance that complements other methods in use. In addition, it may suggest other applications for organizations that deal with risk and return issues in evaluating and managing their performance.

Hedge fund assets traditionally have low correlations with stock and bond markets and offer protection in turbulent markets (Amenc, Bied and Martellini (2003)). Much recent debate has centered on how to measure and evaluate the performance of hedge funds. Comparing hedge funds to standard market indices could be erroneous since hedge funds are viewed as an alternative asset class and possess different characteristics than traditional stock and bond funds. In the literature we frequently observe hedge fund rankings displayed using measures such as the Sharpe ratio, but this could pose problems due to the option-like returns that hedge funds generate (see Fung and Hsieh (1997)). In short, the Sharpe ratio is a risk-adjusted measure or a direct measure of reward-to-risk. It is calculated by subtracting the risk-free rate (U.S. Treasury bill rate) from the returns of a portfolio and dividing it by the standard deviation of the returns. (Standard deviation of returns is widely used as a measure of investment return risk.) Therefore, the higher the Sharpe ratio, the better the portfolio's risk-adjusted performance.

Including hedge funds that use leverage, short selling and other derivatives strategies in an investment portfolio can provide valuable protection against declines in portfolio value as markets gyrate. Nevertheless, the inclusion of hedge funds in investor portfolios further requires more accurate methodologies to handle the skewed distribution of returns they generate. Particularly, hedge fund manager selection requires accurate and objective appraisal of both risk and reward especially when dealing with skewed returns of hedge funds.

Several authors have used multifactor models to examine hedge fund performance (Edwards and Caglayan (2001); Gregoriou, Rouah and Sedzro (2002)). However, as Brealey and Kaplanis (2001) note, there are problems inherent in using these types of traditional approaches in a world of non-normal returns. For example, because of their dynamic trading strategies, hedge funds do not have stable exposure to market factors over time because they use short-selling, leverage and derivatives strategies resulting in low predictive powers of models. Furthermore, hedge funds are known as absolute return vehicles and their aim is to provide superior performance with low volatility in both bull and bear markets as opposed to comparing their relative performance to traditional market indices. It is not unusual that models based on indices do not work well for the non-normality (or skewness) of hedge fund returns. Due to their non-normal characteristics, it is difficult to find appropriate benchmarks, and in some cases traditional benchmarks such as the S&P 500 Index and the Morgan Stanley World Index have been used to compare hedge fund returns resulting in low statistical R-squared values (Gregoriou, Rouah and Sedzro (2002); Edwards and Caglayan (2001)).

Here we compare hedge fund performance using DEA as an alternative measure to the Sharpe ratio. The main advantage of DEA lies in its versatility because of the different units of measure as well as multiple input and outputs can be used.

DEA permits us to appraise and rank hedge funds in a risk-return framework without using indices, which is a key advantage because of the absence of hedge fund performance indices. The power of DEA is in its ability to deal with several inputs and outputs while not demanding a precise relation between input and output variables. Another potential advantage that will be more apparent in the application to hedge funds is that DEA considers more characteristics of the hedge fund via the selection of the multiple input and output measures than currently used measures and consequently provides a more sensitive and comprehensive measure of performance. Use the DEA reference set information, the hedge funds themselves are used as benchmarks in the absence of hedge fund performance indices.

Having an alternative performance measure like DEA is potentially of great value in enabling investors to pinpoint the reasons behind a fund's poor performance and it can be used as an indicator of performance. For institutional investors considering using hedge funds as downside protection in bear markets, it is critical that a performance measure provide not only a precise appraisal of the fund's performance, but also an idea of the method management uses to control risk with respect to certain criteria (variables such as inputs and outputs). Using DEA can provide investors with a useful tool for ranking hedge funds by self appraisal and peer group appraisal.

The balance of the chapter is organized as follows. Section 11.2 presents some background information on the hedge funds. Section 11.3 discusses the data and the method. Section 11.4 presents an application to a set of 634 hedge funds within 8 fund classifications. Section 11.5 concludes.

11.2. BACKGROUND INFORMATION

Hedge funds have frequently been referred to as funds providing "absolute returns" given that their objectives are to offer positive returns irrespective of market conditions whilst not being compared to any benchmarks. It is commonly known that alternative performance evaluation techniques are essential in calculating the risk exposure characteristics of hedge funds.

Performance measurement is an essential tool for investors and has recently become the central issue for understanding the behavior of hedge funds, especially in bear markets. Many investors' insight of hedge fund performance and the fund manager skill is frequently compared to traditional benchmarks such as the S&P 500 and the Morgan Stanley Capital International World index (MSCI World). However, this comparison is more valid for mutual funds than for hedge funds because mutual funds have more normal distributions.

Research has also indicated that there exists tracking error among various hedge fund indices, where some indices are equally weighted while others are value weighted and this could possibly affect the evaluation of performance (McCarthy and Spurgin (1998)). An equally weighted index which uses all the average returns of the hedge funds that make up the index is preferred to a value weighted index which uses assets under management because it reproduces diversification intended to track these types of indices.

Hedge funds have different return characteristics than mutual funds and using standard performance measures to evaluate various hedge fund strategies could be misleading (Fung and Hsieh (2000)). Due to their asymmetric or skewed returns and dynamic trading strategies, the use of risk-adjusted measures, such as the traditional Sharpe ratio, is considered unsuitable for hedge funds and the better suited for the normally distributed returns of mutual funds.

It is difficult to identify factors that drive hedge fund returns, unlike the factors that have been proven effective for mutual funds. Investors and analysts placing too much faith in these models are therefore at risk of being misled, by biased alphas (Schneeweis and Spurgin (1999)). However, the underlying question still remains with reference to which benchmarks would be appropriate for each hedge fund strategy, given that index models could no longer be suitable.

Recent studies, such as Edwards and Caglayan (2001) investigate hedge fund alphas (the ability of the manager to outperform his or her respective benchmark) or manager skill using multifactor models, while Liang (2000) examines survivorship bias of hedge funds or the tendency for returns to be overstated due to the exclusion of dead funds. (Funds that have not reported for three consecutive months are considered dead by the database vendors. However, a small number of dead funds may still be operating and do not want to attract new capital from investors. Therefore, hedge fund managers refrain from sending their monthly returns to database vendors and in many cases this could be a sign that the fund has attained capacity constraints and does not want to advertise itself anymore. Agarwal and Naik (2000) find significant quarterly performance persistence (winners continue to be winners) in hedge funds, while Edwards and Caglayan (2001) observe performance persistence for winners and losers. On the other hand, Brow, Goetzmann and Ibbotson (1999), Peskin et al. (2000), and Ackermann, McEnally and Ravenscraft (1999) uncover slight significant performance persistence, relative to traditional asset classes. The way performance is measured could be the consequence of the divergent results therefore excess return could display performance persistence when in fact it is inexistent.

Amenc, Bied and Martellini (2003) using multifactor regression models partially succeed in explaining the predictability of hedge fund returns for six out of nine hedge fund styles but with low R-squared values ranging from 15.7 per cent to 53.4 per cent. Regardless of the ability of existing and frequently used models to explain hedge fund returns, the dynamic trading strategies and skewed returns remain a serious matter in hedge fund performance literature. Further investigation is warranted to examine this problem by using other innovative methods, possibly DEA.

Style analysis is a method used to determine the exposure of the portfolio's various investments (or asset allocation) and how they account for the variability in the returns of stock and bond portfolios. For example, in the mutual fund industry style examples could be "value or growth". If the asset classes are outlined, then the style of a particular fund can be established by its exposure to the different asset classes. Fung and Hsieh (1997) and Liang (1999) apply Sharpe's factor "style" analysis to hedge funds and find that it does a poor job to explain hedge fund returns – unlike what Sharpe (1992) observes for mutual funds. They attribute the low *R*-squared values to the dynamic strategies of hedge funds. Despite its frequent use, the main drawback of Sharpe's style analysis assumes that the exposure to the individual styles do not vary through time; its main drawback. LHabitant (2002) also argues that returns based style performs poorly for hedge funds as a result of their various investment strategies, especially since style analysis calls for consistency during the investigation period.

Agarwal and Naik (2000) apply mean-variance analysis to show that portfolios consisting of passive asset classes (passive investment in equities and bonds) mixed with non-directional hedge funds, provide a better risk-return tradeoff than portfolios with passive asset classes only at the expense of increased negative skewness. This implies that negative skewness is an unwanted feature because a large negative returns or loss can wipe out months of compounding positive returns. Agarwal and Naik (2000) define hedge funds whose returns exhibit low correlation with market indices as having "non-directional" strategies, and those with high correlation as having "directional" strategies. Some authors are beginning to apply longitudinal analyses to better describe temporal features of hedge fund performance. Brown, Goetzmann and Park (2001) apply survival analysis to estimate the lifetimes of hedge funds and find these are affected by factors such as their size, their performance and their redemption period.

Investors relying strictly on using volatility as a risk measure for hedge funds could prove inadequate due to their non-normal returns, thereby requiring more appropriate measures, such as skewness and kurtosis. Typically mutual funds have normal distributions but hedge funds suffer from skewed and kurtotic returns and traditional risk-adjusted measures such as the Sharpe ratio do not work to assess their performance. This implies that the Sharpe ratio will tend to overestimate the real performance of a hedge fund.

Furthermore, traditional Sharpe ratios will usually overestimate and miscalculate hedge fund performance, given that negative skewness and excess kurtosis are not considered by this risk-adjusted measure (Brooks and Kat, 2001).[1]

Using hedge fund indices to examine performance persistence could also be a drawback, since they are rebalanced and cannot properly reproduce the same composition during an entire examination period. Consequently persistence could be erroneously estimated. DEA allows us to bypass the use of troublesome benchmarks and potentially develop a more comprehensive performance measure.

11.3. DATA AND METHODS

We received selected hedge fund data from Burlington Hall Asset Management made available by Zurich Capital Markets (ZCM) database provided by the Center for International Securities and Derivatives Markets (www.cisdm.org) based at the University of

[1] Non-normality implies that traditional mean-variance analysis (MVA) is not applicable to hedge funds, because of their skewed and kurtotic distribution. The Capital Asset Pricing Model (CAPM) therefore is not appropriate for evaluating hedge funds because the variance and return do not follow accepted theoretical foundations. As investments, hedge funds display that their low variance provides greater returns and their high variance provides lower returns than what the CAPM presumes.

Massachusetts (Amherst, Massachusetts). We examine eight hedge fund styles or classifications during the 1997-2001 and 1999-2001 periods. Two periods are used to observe if the extreme market event of August 1998 had any impact on various classifications. The short sellers and the long only classifications were eliminated since they only contained a handful of funds and are deemed not sufficient for the analysis. The database provider advised us that using a longer time frame, for example, a 7- or 10- ten-year examination period would have resulted in significantly fewer funds for the study. Our data set consists of monthly net returns, whereby both management and performance fees have already been subtracted by the hedge funds and forwarded to ZCM. We do not examine defunct hedge funds.

Modern portfolio theory measures the total risk of a portfolio by using the variance of the returns. But this method does not separate the upside risk, which investors seek, from the downside returns they want to avoid. Variance is not typically a good method for measuring risk, but semi-variance is frequently used and accepted in the investment area to measure downside risk. Returns above the mean can hardly be regarded as risky by investors, but the variance below the mean provides more information during extreme market events which confirms that investors worry more about underperformance than overperformance (Markowitz (1991))[2].

Furthermore, the mean and standard deviations of hedge fund returns could be misleading and higher moments such as skewness and kurtosis will provide a more accurate picture (Fung and Hsieh, (1997)). Since the investor is faced with non-normal distributions when investing in hedge funds, it is not correct only to use the standard deviation as the sole measure of risk. The investor must examine the degree of symmetry the distribution (skewness) has and the probability of extreme positive or negative returns (kurtosis). The introduction of skewness in the inputs and outputs will present some signaling assessment of each hedge fund classification. To correctly assess hedge fund appraisal, skewness does not penalize hedge funds by the upside potential returns. Although hedge funds attempt to maximize returns and minimize risk, this comes as a trade-off, whereby, adding hedge funds to traditional investment portfolios will likely result in high kurtosis and increased negative skewness which are the drawbacks of this alternative asset class. Moreover, hedge fund returns have fat tails resulting in a greater number of extreme events than one would normally anticipate (Fung and Hsieh (2000)). This

[2] Extreme market events include the following: the Asian currency crisis of 1997, the Russian ruble crisis of 1998, and the September 11, 2001 terrorist attacks.

implies that hedge fund returns are inclined to be negatively skewed (fat tails) which consist of positive returns but there are a few instances of extreme losses or a small chance of a big loss.

The inputs and outputs must correspond to the activities of hedge funds for the analysis to make sense. We use six variables in a risk-return framework, three for inputs and three for outputs.

The inputs are: 1) lower mean monthly semi-skewness, 2) lower mean monthly semi-variance, and 3) mean monthly lower return. The outputs are: 1) upper mean monthly semi-skewness, 2) upper mean monthly semi-variance, and 3) mean monthly upper return. The numerical value of return outputs is the value-added of each hedge fund and the 30-day U.S. T-bill rate is subtracted from the monthly net returns. Lower mean returns simply show the average returns below zero while upper mean returns are the average returns above zero. These measures are chosen because higher output values and smaller input values usually indicate better fund performance. Other possible alternative inputs such as percentage of monthly returns below zero and amount of leverage can be used. For outputs percentage of monthly returns above zero number and compound return can be considered as alternate or added outputs in future research.

The data were aggregated into separate DEA runs for the 3-year (1999-2001) and 5-year (1997-2001) periods for each classification. Both examination periods contain the same funds in each classification enabling us to see whether the rankings would differ and if several funds would be efficient in both periods.

Since hedge funds vary their leverage at different times to magnify returns, we employ the VRS model (variable returns to scale) to identify the efficient and inefficient funds.

11.4. RESULTS

Table 11-1 displays the DEA results in terms of the number of efficient and inefficient funds for both examination periods of 1997-2001 and 1999-2001. The results indicate that a great majority of funds (over 80%) are inefficient in a risk-return framework according to the inputs and outputs we use. The reason can possibly be attributed to the various extreme market events such as the Asian currency crisis of October 1997 and the Russian ruble crisis of August 1998 yielding increased volatility in global stock and bond markets.

Tables 11-2 and 11-3 display descriptive statistics for each hedge fund classification. We find that all efficient funds in the 1997-2001 period, except the global emerging funds, display positive skewness, whereas a majority of the inefficient funds in the 1997-2001 period exhibit negative skewness. In addition, inefficient funds exhibit lower mean monthly returns and negative skewness compared with the efficient funds. This suggests that the funds classified as inefficient did perform more poorly than the efficient funds, as would be anticipated. The reasons for this could be that the global macro and sector categories had no extreme positive or negative returns during the 1997-2001 period, which can either be attributed to the use of strategies without leverage, or the use of convex strategies (buying index options) and also hedge of downside risk (buying options). In other words, hedge funds can be convex on the upside by being protected on the downside (call payoff) and they pay the premium by selling the upside (near-at-the-money) and buying options on the extreme negative to reduce volatility and obtain positive skewness. The effect is due to the extreme market event of August 1998 which caused negative skewness for a large majority of inefficient funds. To properly assess the performance of hedge funds, the length of the examination period is not important, but rather the time series of each hedge fund classification must be long enough to include at least one extreme negative market event, as is the case during the 1997-2001 period.

Table 11-1. Number of Efficient and Inefficient Hedge Funds 1997-2001 and 1999-2001

Classification	Efficient	Inefficient	Total
1997-2001			
Funds of Hedge Funds	10 (6%)	158 (94%)	168
Sector	9 (27%)	24 (73%)	33
Global Macro	7 (29%)	17 (71%)	24
Global Emerging	6 (20%)	24 (80%)	30
Global Established	9 (7%)	124 (93%)	133
Event Driven	20 (27%)	53 (73%)	73
Global International	5 (22%)	18 (78%)	23
Market Neutral	5 (4%)	125 (96%)	130
Total	71 (12%)	543 (88%)	614
1999-2001			
Funds of Hedge Funds	20 (12%)	148 (88%)	168
Sector	11 (33%)	22 (67%)	33
Global Macro	8 (33%)	16 (67%)	24
Global Emerging	12 (40%)	18 (60%)	30
Global Established	18 (14%)	115 (86%)	133
Event Driven	21 (29%)	52 (71%)	73
Global International	8 (35%)	15 (65%)	23
Market Neutral	9 (7%)	121 (93%)	130
Total	107 (17%)	507 (83%)	614

Table 11-2. Monthly Statistics of Efficient and Inefficient Funds 1997-2001

	A	B	C	D	E	F	G	H	I	J	K
Efficient Hedge Funds											
Funds of Hedge Funds	1.53	-48.98	48.67	4.22	1.14	6.15	-6.53	-8.28	1.93	2.09	436.98
Event Driven	1.26	-54.29	88.47	5.42	0.65	3.37	-10.87	-11.34	0.94	1.26	60.64
Market Neutral	2.17	-17.2	72.25	5.47	1.92	9.73	1.33	-10.55	0.9	4.95	484.83
Global Macro	1.76	-30.01	46.75	6.33	1.02	3.51	-10.44	-12.97	0.4	0.96	115.16
Global Int'l	0.82	-46.19	33.75	7.66	0.49	1.34	-15.01	-17	0.22	0.42	22.32
Global Emerging	2.76	-63.79	61.78	14.48	-0.27	2.69	-39.4	-30.92	0.18	0.71	36.57
Global Established	2.08	-58.59	85.8	7.78	1.12	4.76	-11.62	-16.03	0.44	1.21	158.03
Sector	2.18	-38.35	62.15	7.79	0.31	2.82	-12.9	-15.94	1.12	1.16	106.31
Inefficient Hedge Funds											
Funds of Hedge Funds	0.93	-28	21.21	2.86	-0.4	4.49	-7.09	-5.73	0.6	0.86	124.06
Event Driven	0.91	-41.65	49.4	4.27	-0.47	4.7	-11.05	-9.02	0.45	0.88	179.24
Market Neutral	1.1	-48.73	97.61	3.09	-0.49	5.61	-6.33	-6.08	0.63	1.18	341.98
Global Macro	0.99	-53.24	25.73	4.94	0.07	1.95	-10.5	-11.65	0.16	0.28	27.15
Global Int'l	0.78	-44.51	29.51	5.97	-0.42	2.17	-16.53	-13.1	0.11	0.13	38.62
Global Emerging	0.81	-73.25	56.11	8.49	-0.61	4.4	-25.47	-18.94	0.06	0.08	162.82
Global Established	1.29	-48.38	47	5.91	-0.02	2.27	-13.79	-12.46	0.25	0.63	37.75
Sector	1.5	-31.63	42.49	7.4	0.2	2.6	-15.71	-15.81	0.17	0.55	92.36

Column A: Mean (%), Column B: Minimum one period retrun, Column C: Maximum one period return, Column D: Standard deviation, Column E: Skewness, Column F: Excess Kurtosis, Column G: Modified VaR 95%, Column H: Normal VaR, Column I: Modified Sharpe ratio, Column J: Sharpe ratio, Column K: Jarque-Bera

Table 11-3. Monthly Statistics of Efficient and Inefficient Funds 1999-2001

	A	B	C	D	E	F	G	H	I	J	K
Efficient Hedge Funds											
Funds of Hedge Funds	1.48	-21.3	48.67	3.29	0.85	2.75	-3.35	-6.17	2.02	2.19	70.18
Event Driven	1.17	-54.3	88.47	5.61	1.1	3.36	-8.01	-11.9	1.02	1.21	47.54
Market Neutral	2.16	-15.9	72.25	4.29	1.45	5.36	3	-7.81	2.13	4.17	124.06
Global Macro	1.52	-30.0	44.98	6.5	1.1	3.68	-9.97	-13.6	0.25	0.86	60.21
Global Int'l	1.55	-46.2	33.75	7.8	0.7	1.15	-13.3	-16.6	0.28	0.42	10.91
Global Emerging	4.36	-27.7	61.78	12.1	0.93	1.74	-16.6	-23.8	0.55	1.11	14.32
Global Established	1.81	-58.6	85.8	7.77	0.98	2.59	-11.3	-16.3	0.94	1.13	33.79
Sector	2.35	-38.4	62.15	7.66	0.64	1.29	-10.4	-15.5	0.47	1.19	19.94
Inefficient Hedge Funds											
Funds of Hedge Funds	1.01	-22.9	20.2	2.6	0.57	1.74	-3.98	-5.05	1.15	1.22	15.66
Event Driven	0.92	-41.7	49.4	3.68	0.14	1.38	-6.93	-7.65	1.1	1.28	10.14
Market Neutral	1.1	-44.8	97.61	2.58	0.16	2.46	-4.54	-5.53	1.18	1.42	41.21
Global Macro	0.86	-53.2	61.32	0.45	0.27	1.07	-9.2	-9.51	0.22	0.3	6.78
Global Int'l	0.94	-30.8	60.15	5.58	0.06	0.5	-12.1	-12.1	0.18	0.3	3.15
Global Emerging	1.87	-31.8	56.11	7.35	0.36	0.89	-13.2	-15.2	0.34	0.77	0.27
Global Established	1.03	-48.4	87.12	5.91	0.36	1.47	-11.2	-12.7	0.22	0.43	0.08
Sector	1.15	-38.4	66.51	7.25	0.46	1.01	-13.1	-15.7	0.16	0.36	8.72

Column A: Mean (%), Column B: Minimum one period retrun, Column C: Maximum one period return, Column D: Standard deviation, Column E: Skewness, Column F: Excess Kurtosis, Column G: Modified VaR 95%, Column H: Normal VaR, Column I: Modified Sharpe ratio, Column J: Sharpe ratio, Column K: Jarque-Bera

Tables 11-2 and 11-3 provide the Sharpe ratio and the modified Sharpe ratio. For risk-adjusted performance, the traditional Sharpe ratio is not applicable for non-normal returns, thus we use instead the modified Sharpe ratio. Value-at-Risk (VaR) is need in calculating the modified Sharpe ratio. The normal VaR is a method used to estimate a fund's maximum expected loss based on statistical analysis of historical returns and volatility as long as normal returns exist. Since hedge funds have skewed returns we compare the normal VaR with the more appropriate modified VaR for all hedge fund classifications and both periods. The comparison will provide us with a more precise picture because the modified VaR takes into account skewness and kurtosis, whereas normal VaR considers only the mean and standard deviation. The modified VaR allows the calculation of VaR for distributions with either positive or negative skewness as well as positive excess kurtosis (or more commonly known as fat tails). We do not discuss the derivation of the modified VaR and the reader is directed to Favre and Galeano (2002). Tables 11-2 and 11-3 also provides the Jarque-Bera statistic test of non-normality which considers both skewness and excess kurtosis to determine whether all classifications in both examination periods exhibit non-normal distribution of returns. The Jarque-Bera statistic is particularly useful for a large number of monthly returns but not practical for small samples. A Jarque-Bera result greater than 6 indicates the distribution is non-normal.

In Table 11-2 the normal Sharpe ratio is higher than the modified Sharpe ratio because the modified frontier computed in a mean modified VaR is often shifted slightly downwards and to the right of the normal frontier (if skewness is negative and excess kurtosis is positive). When comparing the standard deviations in Tables 11-2 and 11-3 we find that both efficient and inefficient funds in the 1999-2001 period had lower standard deviations than the 1997-2001 periods because of the extreme market event of August 1998. This implies that even though hedge funds typically have low correlations to stock and bond markets in extreme market events, efficient and inefficient funds are affected by market volatility and tend to become more highly correlated to stock markets in severe market corrections. In both periods (Tables 11-2 and 11-3) the results display that efficient funds have higher mean returns, positive skewness (more so than in Table 11-3) for efficient funds during the 1999-2001 period. The skewness is positive for both efficient and inefficient funds during the 1999-2001 period but efficient funds possess large positive skewness than inefficient funds. This can be attributed to the absence of an extreme market event. In the 1997-2001 period where an extreme market event (the fall of the Ruble) in August 1998, efficient funds have positive skewness whereas inefficient funds have negative skewness, which is what investors dislike. Positive skewness, implies that there is a greater probability of a fund experiencing a small loss which would be counterbalanced by a small probability of achieving a large gain. Negative skewness on the other hand implies a small probability of a fund experiencing a large loss would be counterbalanced by a large probability of achieving a small gain. The modified Sharpe ratio is more accurate than the traditional Sharpe ratio because the former ratio factors in skewness and kurtosis which is more appropriate for evaluating hedge funds as a risk-adjusted performance measure. In addition, excess kurtosis is higher for efficient funds in both examination periods, implies that returns are closer to the mean, on average, for efficient funds than inefficient funds and that efficient funds may have extraordinary large gains or large losses. In other words, a higher excess kurtosis implies that the variance in returns is caused by the sporadic extreme deviations.

Investors prefer to reduce the extreme negative events and favor positive as opposed to negative skewness (implying the left tail is fatter than the right tail), since the underlying motivation for hedge funds is their ability to obtain positive returns in flat or down markets. Furthermore, hedge funds advertise (sell) extreme risk which gives negative skewness and positive kurtosis which can be compared to a short put option (Agarwal and Naik (2000)). Adding hedge funds to a traditional stock and bond portfolio to obtain higher risk-adjusted returns and lower volatility will result in a trade-off between negative skewness and diversification of the portfolio.

Hedge fund returns do not follow normal distributions because their returns are asymmetrical and display fat tails, a finding validated in both Tables 11-2 and 11-3, whereby the non-directional strategies[3] display fatter tails (excess kurtosis) than the directional strategies. As well, the non-directional strategies possess lower volatility than directional ones, a fact that is widely known.

Obviously the results in Table 11-2 indicate somewhat high excess kurtosis (fat tails) in the non-directional strategies (event driven and market neutral). More frequently excess kurtosis is calculated as kurtosis minus 3 to simplify the explanation. A normal distribution has a kurtosis of 3 and an excess kurtosis of zero implies a normal distribution. Whereas, an excess kurtosis greater than zero signifies a

[3] The non-directional strategies include market neutral and event driven funds, whereas the directional strategies include the rest.

high probability of big gains or losses. The greater the positive excess kurtosis the more the distribution will be peaked or leptokurtic. This implies that there are more returns close to the mean with more frequent large positive or large negative returns than a normal distribution of returns. This signifies that there is a high probability that extreme market events will occur making this distribution spiked when compared to the normal distribution. The possible reason for this occurrence is that the non-directional classification possesses payoffs like short option strategies whereas, the other directional strategies possess long only option strategies. Therefore, hedge funds usually possess more positive excess kurtosis (fat tails) than traditional normal distributions. A fat tailed distribution will generally have a greater number of recurrent extreme (larger or smaller) observations than a typical normal distribution, a finding displayed in the non-directional classification and funds of hedge funds where volatility is commonly known to be the lowest. The non-directional strategies attempt to take advantage of irregularities in stock and bond markets and perform well during stable market conditions.

However, the inefficient global macro fund classification displays positive skewness during the 1997-2001 period, because this style requires movement of global markets to profit from major trends and destabilizing market conditions. When negative skewness is present in the data, it implies that the payoffs of hedge funds are exposed to the downside (decline in value) more than normally distributed funds. The number of funds with negative skewness is not necessarily good or bad, it merely implies that investors familiar with risk management will be aware of a decrease in expected return will eventually occur to bear this negative skewness.

Furthermore, we notice that non-directional funds (event driven and market neutral) have a low standard deviation and high excess kurtosis whereas, for directional funds (or sometimes called market timing funds, for example, global macro, global international, global emerging, global established and sector) we observe a high standard deviation with low excess kurtosis, given that they have a greater exposure to market risk than non-directional funds.

Since hedge funds use dynamic strategies and produce non-linear payoffs, we find that in Table 11-2 the modified VaR is the highest for the market neutral classification. A high modified VaR implies the modified VaR is near to zero, therefore, a high modified Sharpe ratio is due to a modified VaR near zero. As we approach a modified VaR of zero, the modified Sharpe increases exponentially. In other words, modified VaR penalizes funds with extreme negative returns as the modified VaR accounts for negative skewness, disliked by investors, and for positive excess kurtosis disliked by investors. The modified VaR measures the risk of losing 5 per cent or more 1 per cent of the time. On the other hand, the global emerging classification has the highest modified VaR during the 1997-2001 period. The difference between the normal and modified VaR comes from the asymmetries in the hedge fund returns distribution (skewness) and from the positive or negative extreme returns (kurtosis). Comparing both normal and modified VaR will illustrate the impact of neglecting extreme market returns of the measure used in a normal VaR. Non-normal distributions are due to negative skewness (concave payoffs due to premium selling) and /or to excess kurtosis due to extreme market events (liquidity/event risk).

When we examine the Jarque-Bera statistic during both examination periods, we observe that Table 11-2 specifies that during the 1997-2001 period all distributions are non-normal with a greater amount of non-normality for non-directional funds as a result of the extreme market event of August 1998. However, during the 1999-2001 period, extreme market events were inexistent thereby producing a lower degree of non-normality with all classifications. Furthermore, of the inefficient hedge funds during the 1999-2001 period displayed in Table 11-3, three directional classifications (global international, global emerging and global established) have non-normal distributions.

We also notice that in both Tables 11-2 and 11-3 the standard deviations are higher for the efficient funds in both periods, but using a one tailed t-test the p-value is only significant for the 1999-2001 period (p=.040). Because we suspected a priori that efficient funds would have higher mean monthly returns and higher skewness when compared to non-efficient funds, we use a one-tailed t-test. The results indicate that mean monthly returns and skewness for efficient funds are higher than non-efficient funds during both the 1997-2001 and the 1999-2001 periods at the 1 per cent level with p-values of p=.004 and p=.013 respectively for returns and p=.0005 and p<.0001 respectively for skewness.

In Table 11-3 we discover that all efficient and inefficient funds exhibit positive skewness and can be explained by the lack of extreme market events during the 1999-2001 period. The sole directional classification (global macro) benefited from positive skewness during this period owing to its strategy that is based on global economic indicators, as well as to political and macroeconomic views of different countries. Therefore, the global macro classification during the 1999-2001 period could in fact, have had little exposure to market events as indicated by the lower modified VaR when compared to the 1997-2001 period. As well, after August 1998 a great many global macro hedge funds closed due to their use of excess leverage.

Furthermore, the Basle Committee on Banking Supervision and the Sound practices for Hedge Fund Managers report both have managed to impose on the remaining global macro hedge funds in operation more preventative steps in terms of using excess leverage. Recently, Gregoriou, Rouah and Sedzro (2002) observed that the global macro classification experienced the second lowest survival time with a half-life of 3.59 years during the 1990-2001 period.

Table 11-4. **Champion Hedge Funds 1997-2001 and 1999-2001**

Champion Hedge Funds 1997-2001		
HEDGE FUND	REFERENCE SET	CLASSIFICATION
HALCYON SPECIAL SITUATIONS	44	Event Driven
PRIME ADVISORS FUND LTD	157	Fund of Hedge Funds
VAN ECK GLOBAL	23	Global Emerging
SEMINOLE CAPITAL PARTNERS	89	Global Established
CAXTON GAM	14	Global Macro
ARTIC HEDGE FUND	20	Global International
KCM BIOMEDICAL	25	Sector
ATLANTIS CAPITAL MANAGEMENT	94	Market Neutral
Champion Hedge Funds 1999-2001		
HEDGE FUND	REFERENCE SET	CLASSIFICATION
TWIN SECURITIES	49	Event Driven
LAFAYETTE EUROPE FUND	122	Fund of Hedge Funds
ASHMORE RUSSIAN DEBT	12	Global Emerging
CIRCLE T PARTNERS	37	Global Established
CAXTON GAM	14	Global Macro
ARTIC HEDGE FUND	20	Global International
KCM BIOMEDICAL	19	Sector
CLARION OFFSHORE FUND	88	Market Neutral

In Table 11-4 identifies the "champion' hedge funds"[4]. The score implies the number of times an efficient fund has been part of an inefficient hedge fund's reference set as a result of the DEA efficiency

[4] These champion funds are located by the frequency of appearance in the ERS of inefficient funds.

analysis. As the frequency of a hedge fund appearing in a reference set increases, the likelihood of the fund being a good performer increases. The efficient hedge fund appearing in the most reference sets can be considered the overall "champion" and can help inefficient funds learn from their superior management and investment practices. As well, the reference set of a hedge fund can shed some light as to why a fund is performing poorly and display potential improvements in its weak areas. The 1997-2001 champion is the Prime Advisors Fund; the 1999-2001 champion is the Lafayette Europe Fund, both funds of hedge funds. These funds achieved the greatest compound return with the least amount volatility; a common requirement by investors whereby they can add these funds to traditional stock and bond portfolios to reduce the overall portfolio risk and enhance performance.

Finally, if we look at the individual funds, we observe that in all classifications there are fewer efficient funds in the 5-year as opposed to the 3-year period. The reason is that funds during the 1997-2001 period could have experienced a tremendous amount of increased risk owing to the Asian currency crisis of 1997 followed by the Russian ruble crisis of August 1998. It is also surprising to see that the fund of hedge funds, and market neutral classifications contain the least number of efficient funds in both examination periods when compared to the total number of funds (as a percentage) in their respective categories. The abnormality could be partly due to the high kurtosis (fat tails) in each of the above classifications exposing them to extreme market events.

The large majority of efficient funds during the 5-year period are also efficient in the 3-year period, possibly providing an indication that some hedge funds are able to control for risk with a greater amount of accuracy than other funds. The large majority of funds that were efficient in both periods had higher cross efficiency scores in the 5-year period.

11.5. CONCLUSIONS

This chapter demonstrates the potential for using DEA as an alternative selection tool to assist pension funds, institutional investors, FOF managers and high-net worth individuals in selecting efficient hedge funds. We believe DEA is a valuable complement to other risk-adjusted measures because it presents a more complete picture of hedge fund performance appraisal both because it uses the powerful DEA methodology and because this enables one to incorporate more key dimensions about fund performance in the assessment of the funds. . Hedge fund performance evaluation using DEA is potentially valuable because it allows an investor to more accurately identify superior performing funds compared to conventional risk-measurement techniques, which can be misleading.

Using DEA can help investors to re-examine performance measurement and validate the selection process of hedge funds especially when there exists over 7,000 funds in the universe. New methods should be used to add consistency to the manager selection process. The most attractive funds to select based on the DEA analysis as reflected in Table 11-4 are the Prima Advisors Fund and the Lafayette Europe Fund, both funds of hedge funds. As the methods of selecting and measuring inputs and outputs for DEA assessments of fund performance, DEA may prove valuable as an independent performance assessment tool.

Future research using other DEA models could examine the efficiency managed futures classifications or even further examine the trading efficiency of these money managers responsible for trading futures. It would also be interesting to measure the efficiency of various hedge fund indices from database vendors, such as Hedge Fund Research (HFR), EACM, ALTVEST, and TASS.

Finally, other DEA approaches can also be used to characterize the performance of the hedge funds. For example, one can use the DEA-based benchmark models to compare the performance of hedge funds in different groups. One can also use the recent development of super-efficiency in Chen (2004; 2005) to fully rank the hedge fund performance.

This chapter is based upon Gregoriou[5], G.N., Sedzro[6], K. and Zhu, J., Hedge fund performance appraisal using data envelopment analysis, European Journal of Operational Research, Vol. 164, Issue 2 (2005), 555-571, with permission from Elsevier.

[5] Gregoriou is Associate Professor of Finance at State University of New York (Plattsburgh).

[6] Sedzro is Professor of Finance at University of Quebec at Montreal.

REFERENCES

1. Ackermann, C., R. McEnally and D. Ravenscraft (1999), "The Performance of Hedge Funds: Risk, Return and Incentives," *Journal of Finance* 54 (3), 833-874.
2. Agarwal, V. And N.Y. Naik (2000), "On Taking the Alternative Route: Risks, Rewards, and Performance Persistence of Hedge Funds," *Journal of Alternative Investments* 4 (2), 6-23.
3. Ali, A. I., C.S. Lerme and L.M. Seiford (1995), "Components of Efficiency Evaluation in Data Envelopment Analysis," *European Journal of Operational Research*, Vol. 80, 462-473.
4. Amenc, N., S.E., Bied and L. Martellini (2003), "'Predictability in Hedge Fund Returns," *Financial Analysts Journal*, vol. 59, No. 5, 32-46.
5. Andersen, P. and N.C. Petersen (1993), "A Procedure for Ranking Efficient Units in Data Envelopment Analysis," *Management Science*, Vol. 39, 1261-1264.
6. Athanassopoulos, A. D., and D. Giokas (2000), "The Use of DEA in Banking Institutions: Evidence from the Commercial Bank of Greece," *Interfaces,* Vol. 30, No. 2, 81-95.
7. Athanassopoulos, A. D. (1997), "Service Quality and Operating Efficiency Synergies for Management Control in the Provision of Financial Services: Evidence from Greek Bank Branches," *European Journal of Operations Research,* Vol. 98, No. 2, 300-313.
8. Bank Technology Report (1992), "Analysis Technique Adresses Branch Productivity," Vol. 21, No. 12 (May), p.2.
9. Banker, R.D., R.F. Cannon and R.D. Strauss (1986), "A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production," *Management Science* 32, 30-44

10. Banker, R.D., A. Charnes, W.W. Cooper, J. Swarts, and D. Thomas (1989), "An Introduction to Data Envelopment Analysis with Some of Its Models and Their Uses," In: J.L. Chan and J.M. Patton., eds. *Research in governmental and nonprofit accounting,* Vol. 5, 125-163.
11. Banker R.D. and R. Morey (1986), "Efficiency Analysis for Exogenously Fixed Inputs and Outputs," *Operations Research,* Vol. 34, 513–21.
12. Berg, S., F.R. Førsund and E.S., Jansen (1990), "Deregulation and Productivity Growth in Norwegian Banking 1980-1998: A Non-parametric Frontier Approach," Bank of Norway, Oslo, Norway.
13. Bergman, D.A. and R. Pantell (May 1986), "The Impact Of Reading A Clinical Study On Treatment Decisions of Physicians And Residents," *J. Med. Ed,* Vol. 61, 380-387.
14. Bertakis, K.D.and J.A. Robbins (1987), "Gatekeeping In Primary Care: A Comparison Of Internal Medicine and Family Practice," *The Journal of Family Practice,* 24(3), 305-309.
15. Bessent, A., W. Bessent, J. Elam, and D. Long (1984), "Educational Productivity Council Employs Management Science Methods to Improve Educational Quality. *Interfaces*, Vol. 14, No. 6, 1-8.
16. Brealey, R.A. and E. Kaplanis (2001), "Hedge Funds and Financial Stability: An analysis of Their Factor Exposures," *International Finance* 4(2), 161-187.
17. Brooks, C. and H. Kat (2001), "The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors, Working paper, University of Reading, ISMA Centre.
18. Brown, S.J., W.N. Goetzmann and R.G. Ibbotson (1999), "Offshore Hedge Funds, Survival and Performance 1989-1995," *The Journal of Business* 72 (1), 91-117.
19. Brown, S. J., W.N. Goetzmann and J. Park (2001), "Careers and Survival: Competition and Risk in the Hedge Fund and CTA Industry," *Journal of Finance* 56 (5), 1869-1886.
20. Burgess, J.F. and P. W. Wilson (1996), "Hospital Ownership Aand Technical Efficiency," 42, No. 1, 110-123.
21. Burgess, J.F. and P. W. Wilson (1998), "Variation In Inefficiency Among U.S. Hospitals", *INFOR* 36, 84-102.
22. Callen, J. (1991), "Data Envelopment Analysis: Partial Survey and Applications for Management Accounting," *Journal of Management Accounting Research*, Vol. 3, 35-56.

23. Charnes, A. and W.W. Cooper, 1961, *Management Models and Industrial Applications of Linear Programming,* 2 vols., John Wiley and Sons, Inc., New York.
24. Charnes, A., W.W. Cooper, B. Golany, L.M. Seiford and J. Stutz (1985), "Foundations of Data Envelopment Analysis for Pareto-Koopman's Efficient Empirical Production Functions," *J. of Econometrics,* Vol. 30, 1-17.
25. Charnes, A., W.W. Cooper, Z.M. Huang, and D.B. Sun (1990), "Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks," *Journal of Econometrics*, Vol. 46, No. 1-2, pp. 73-91.
26. Charnes, A., W.W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, Vol. 2, No. 6, 429-444.
27. Charnes, A., W.W. Cooper and E. Rhodes (1981), "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," *Management Science*, Vol. 27, No. 6, 668-697.
28. Charnes, A., W.W. Cooper, and S. Li (1989), "Using DEA to Evaluate Relative Efficiencies in the Economic Performance of Chinese Cities," *Socio-Economic Planning Sciences*, Vol. 23, 325-344.
29. Chen, Y. (2004), "Ranking Efficient Units in DEA," *OMEGA,* Vol. 32, 213-219.
30. Chen, Y. (2005), "Measuring Super-efficiency in DEA in the Presence of Infeasibility," *European Journal of Operational Research*, Vol. 161, 545-551.
31. Chen, Y. and H.D. Sherman (2004), "A Non-Radial Supper-Efficiency DEA Approach to Assessing NATO Burden Sharing," *Socio-Economic Planning Sciences*, Vol. 38, Issue 4, 307-320.
32. Chen, Y. and J. Zhu (2003), "DEA Models for Identifying Critical Performance Measures," *Annals of Operations Research*, Vol.124, 225-244.
33. Cherkin, D.C., et al (1987), "The Use of Medical Resources By Residency-trained Family Physicians And General Internists," *Medical Care*, 25(6), 455-469.
34. Childs, A.W. and D. Hunter (1972), "Non-Medical Factors Influencing Use Of Diagnostic X-ray By Physicians," *Medical Care*, 10 (4), 323-335.
35. Chilingerian, J.A. (1989), "Investigating Non-medical Factors Associated with the Technical Efficiency of Physicians in The Provision Of Hospital Services: A Pilot Study, Best Paper

Proceedings, "*Annual Meeting Of The Academy of Management, Washington, D.C.*
36. Chilingerian, J., A., M. Glavin and S. Bhalotra (2002), "Using DEA To Profile Cardiac Surgeon Efficiency," *Draft of Technical Report to AHRQ,* Heller School, Brandeis University
37. Chilingerian, J. and H.D. Sherman (1990), "Managing Physician Efficient and Effectiveness in Providing Hospital Services," *Health Service Management Research*, Vol. 3, No. 1, 3-15.
38. Chilingerian, J. and H.D. Sherman (1994), "Evaluating and Marketing Efficient Physicians Toward Competitive Advantage," *Health Care Strategic Management*, Vol. 12, No. 5, 16-19.
39. Chilingerian, J.A.and H.D. Sherman (1997), "DEA And Primary Care Physician Report Cards: Deriving Preferred Practice Cones From Managed Care Service Concepts And Operating Strategies," *Annals of Operations Research*, Vol. 67, 83 – 116.
40. Chilingerian, J. A and Glavin, M. (1994), "Temporary Firms In Community Hospitals: Elements of A Managerial Theory Of Clinical Efficiency" *Medical Care Research and Review,* 51 No.3.
41. Churchill, N.C., W.W. Cooper and V. Govindarajan (1982), "Effects of Audits on Behavior of Medical Professionals Under the Bennett Amendment," *Auditing – A Journal of Practice and Theory*, Vol. 1, No. 2, Winter.
42. Cook, W.D., (2004), "Qualitative Data in DEA",in *Handbook on Data Envelopment Analysis*, Chapter 6, eds W.W. Cooper, L.M. Seiford and J. Zhu, Kluwer Academic Publishers, Boston.
43. Cook, W.D., L.M. Seiford and J. Zhu (2004), "Models for Performance Benchmarking: Measuring the Effect of E-commerce Activities on Banking Performance, *OMEGA*, Vol. 32, Issue 4, 313-322
44. Cook, W.D. and J. Zhu (2006), "Rank Order Data in DEA: A General Framework," *European Journal of Operational Research*, (forthcoming).
45. Cooper, W.W., S. Li, L.M. Seiford and J. Zhu (2004), "Sensitivity Analysis in Data Envelopment Analysis," in *Handbook on Data Envelopment Analysis*, Chapter 3, 75-97, eds W.W. Cooper, L.M. Seiford and J. Zhu, Kluwer Academic Publishers, Boston.
46. Cooper, W.W., K.S. Park, and J. Pastor (1999), "RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA," *Journal of Productivity Analysis*, Vol. 11, 5-42.
47. Cooper, W.W., L.M. Seiford and K.Tone (2000), *Data Envelopment Analysis: A Comprehensive Text with Models,*

Applications, References and DEA-Solver Software, Boston: Kluwer Academic Publishers.

48. Cooper, W.W., L.M. Seiford and J. Zhu (2004), *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers, Boston.
49. Darton H. (1992), "The Measurement of the Inequality of Incomes," *Economic Journal*, Vol. 30, 384-361.
50. Diewert, W.E. (1981), "The economic theory of index numbers: a survey," pp. 163-208 in *Essays in the Theory and Measurement of Consumer Behavior in Honor of Sir Richard Stone*, A. Deaton (ed.), London: Cambridge University Press.
51. Edwards, F. and M.O. Caglayan (2001), "Hedge Fund Performance and Manager Skill," *Journal of Futures Markets* 21 (11), 1003-1028.
52. Eisenberg, J. (1986) "Doctors Decisions And The Cost Of Medical Care," *Health Administration Press,* Ann Arbor, M.
53. Eisenberg, J.and D. Nicklin (1981), "Use Of Diagnostic Services By Physicians In Community Practice," *Medical Care,* 19 (3), 297-309.
54. Färe, R., S. Grosskopf and C.A.K. Lovell (1994), *Production Frontiers*. Cambridge University Press.
55. Favre, L. And J.A. Galeano (2002), "Mean Modified Value-At-Risk Optimization with Hedge Funds," *Journal of Alternative Investments* 5 (2), 21-25.
56. Ferrier, G. D.and V. Valdanis (1996), "Rural Hospital Performance And Its Correlates," *The Journal of Productivity Analysis*, 7, 63-80.
57. Fetter, R.B.and J. Freeman (1986), "Diagnostic Related Groups: Product Line Management With In Hospitals," *Academy of Management Review, XI,* 1, 41-54.
58. Fishbane, M. and B. Starfield (1981), "Child Health Care In the United States," *New England Journal of Medicine,* 305, 552-556.
59. Freeborn, D.K. (1972), "Determinants Of Medical Care Utilization: Physicians' Use Of Laboratory Services," *AJPH,* 62(6), 846-853.
60. Fung, W. and D. Hsieh (1997), "Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds," *Review of Financial Studies* 10 (2), 275-302.
61. Fung, W. and D. Hsieh (2000), "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases," *Journal of Financial and Quantitative Analysis* 35 (3), 291-307.

62. Golany, B. and S. Thore (1997), "The Economic and Social Performance of Nations: Efficiency and Returns to Scale," *Socio-Economic Planning Sciences*, Vol. 31, 191-204.
63. Greenfield, S., E. Nelson, M. Zubkoff, W. Manning, W. Rogers, R. Kravitz, A. Keller, A. Tarlov, J. Ware Jr (March 25, 1992), "Variations In Resource Utilization Among Medical Specialties And Systems of Care," *JAMA,* Vol. 267, No.12: 1624-1630.
64. Gregoriou, G. N., F. Rouah and K. Sedzro (2002), "On the Market Timing of Hedge Fund Managers," *Journal of Wealth Management* 5 (1) 26-38.
65. Hashimoto, A. and H. Ishikawa (1993), "Using DEA to Evaluate the State of Society as Measured by Multiple Social Indicators," *Socio-Economic Planning Sciences*, 27, 257-268.
66. Hollingsworth, B., P.J. Dawson and N. Maniadakis (1999), "Efficiency Measurement Of Health Care: A Review Of Non-parametric Methods And Applications," *Health Care Management Science,* 2, No. 3, 161-172.
67. Iida, J. (1991), "US Bancorp Seeks Savings via Branch-analysis System," *American Banker*, October 8, Vol. 156, No. 195, p.3.
68. InterStudy (1989), "The InterStudy Edge. Excelsior, " *MN: Interstudy*
69. Kerr, C., J. C. Glass, G. M. McCallion and D. G. McKillop(1999), "Best-practice Measures Of Resource Utilization For Hospitals: A Useful Complement In Performance Assessment," Public Administration, 77, No.3, 639-650.
70. Knox, R. (1992)."Specialist Care Seen as Driving Cost," *The Boston Globe,* Wednesday, March 25.
71. Kozmetsky, G. and P. Yue (2005), The Economic Transformation of the United States 1950-2000, West Lafayette, Inc. Purdue University Press.
72. Lewin, A.Y. and J. W. Minton (1986), "Determining Organizational Effectiveness: Another Look And An Agenda For Research," *Management Science.* 32, No. 5: 514-538.
73. LHabitant, S. (2002), "Assessing Market Risk for Hedge Funds and Hedge Fund Portfolios," *The Journal of Risk Finance* 2 (4), 16-32.
74. Liang, B. (1999), "On the Performance of Hedge Funds," *Financial Analysts Journal* 55 (4), 72-85.
75. Lynch, J., Y. Ozcan (1994), "Hospital Closure: An Efficiency Analysis," *Hospital and Health Services Administration,* 39, No. 2, 205-212.

76. Luke, R. D., Y. A. Ozcan (1995), "Local Markets and Systems: Hospital Consolidations In Metropolitan Areas," *Health Services Research*, 30, No. 4, 555-576.
77. Maasoumi, E. and G. Nickelsburg (1998), "Multivariate Measures of Well Being and An Analysis of Inequality," *Journal of Business and Economic Statistics*, 6, 327-334.
78. Markowitz, H. (1991), *Portfolio Selection: Efficient Diversification of Investments.* Blackwell Publishers, London.
79. McCarthy, D. and R. Spurgin (1998), "Comparisons of Alternative Hedge Fund Benchmarks," *The Journal of Alternative Investments* 1(1), 18-28.
80. Miller, Robert H. and Harold S. Luft (Winter 1991), "Perspective Diversity And Transition In Health Insurance Plans," *Health Affairs,* Vol. 10, No. 4, pp. 37 - 47.
81. Noren, J. (1980), "Ambulatory Medical Care: A comparison of Internists And Family-general Practitioners," *New England Journal of Medicine*, 302,11-16.
82. Nunamaker, T. (1983), "Measuring Routine Nursing Service Efficiency: A Comparison of Cost Per Day And Data Envelopment Analysis Models," *Health Services Research,* XVIII (2) Part 1, 183-205.
83. O'Keefe, M. (1994), "Different Perspectives – Cutting Through Benchmarking Jargon," *Bank Systems Technology*, November, p. 53.
84. Ozcan, Y., J. Lynch (1992), "Rural Hospital Closures: An Inquiry Into Efficiency," *Advances in Health Economics and Health Services Research,* 13, 205-224.
85. Perez, E.D. (1992), "Regional Variations In VAMC's Operating Efficiency," *Journal of Medical Systems*, 16, No. 5, 207-213.
86. Peskin, M., M., Urias, S. Anjilvel, and B. Boudreau (2000), *Why Hedge Funds Make Sense*, Quantitative Strategies Morgan Stanley Dean Witter, New York.
87. Pyrrh, P.A., (1973), *Zero-based Budgeting*, John Wiley & Son, New York.
88. Rosenblatt, Roger. (March 25, 1992), "Specialists Or Generalists: On Whom Should We Base The American Health Care System," *JAMA*, Vol. 26, No. 12,: 1665-1666.
89. Rosenthal, E. (1994), "Doctors Slow To Join H.M.O.'s Now Often Find Doors Shut," *The New York Times*, June 25, 1-26.
90. Rouse, P., M. Putterill and D. Ryan, (2002), "Integrated Performance Measurement Design: Insights from an Application in

Aircraft Maintenance," *Management Accounting Research,* Vol. 13, No. 2, 229-248.

91. Rousseau, J. J. and J.H. Semple (1993), "Categorical Outputs In Data Envelopment Analysis," *Management Science,* Vol 39, 392 - 386.

92. Schneeweis, T and R. Spurgin (1999), "Alpha, alpha... Who's Got the Alpha?," *Journal of Alternative Investments* 2(2), 83-87.

93. Schroeder, R.G. (1985), *Operations Management*, McGraw-Hill, New York.

94. Seiford, L.M. and J. Zhu (1988), "Identifying Excesses and Deficits in Chinese Industrial Productivity (1953-1990): A Weighted Data Envelopment Analysis Approach," *OMEGA,* Vol. 26, No. 2, 279-296.

95. Seiford, L.M. and J. Zhu (1988b), "An acceptance system decision rule with data envelopment analysis," *Computers and Operations Research*, 25, 329-332.

96. Seiford, L.M. and J. Zhu (1999a), "An Investigation of Returns to Scale under Data Envelopment Analysis," *OMEGA,* Vol. 27, No. 1, 1-11.

97. Seiford, L.M. and J. Zhu (1999b), "Profitability and Marketability of the Top 55 US Commercial Banks," *Management Science*, Vol. 45, No. 9, 1270-1288.

98. Seiford, L.M. and J. Zhu (1999c), "Infeasibility of Super Efficiency Data Envelopment Analysis Models," *INFOR, Information Systems & Operational Research,* Vol. 37, No. 2, 174-187.

99. Seiford, L.M. and J. Zhu (2003), "Context-dependent Data Envelopment Analysis: Measuring Attractiveness and Progress. *OMEGA*, Vol. 31, Issue 5, 397-480.

100. Sexton, T., A. Leiken, A. Nolan, S. Liss, A. Hogan, R. Silkman (1989), "Evaluating Managerial Efficiency Of Veterans Administration Medical Centers Using Data Envelopment Analysis, " *Medical Care*, 27, No.12, 1175-1188.

101. Sharpe, W.F. (1992), "Asset Allocation: Management Style and Performance Measurement," *Journal of Portfolio Management*, 18 (2), 7-19.

102. Sherman, H.D. (1984), "Hospital Efficiency Measurement And Evaluation, Medical Care," 22, No. 10, 922-928.

103. Sherman, H. David (1988), *Service Organization Productivity Management,* The Society of Management Accountants Of Canada, Hamilton, Ontario, Canada.

104. Sherman, H.D. (1989), *Service Organization Productivity Management – Highlights Edition*, The Society of Management Accountants of Canada, Hamilton, Ontario, Canada.
105. Sherman, H.D. and F. Gold (1985), "Bank Branch Operating Efficiency: Evaluation with Data Envelopment Analysis," *Journal of Banking and Finance*, Vol. 9, No. 2, 297-315.
106. Sherman, H.D. and G. Ladino (1995), "Managing Bank Productivity Using Data Envelopment Analysis (DEA)," *Interfaces* Vol. 25, No. 2, 60-73.
107. Simonson, I. and A. Tversky (1992), "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, Vol. 29, 281-295.
108. Slottje, D.J., G.W. Scully, J.G. Hirschberg, and K.J. Hayes (1991), *Measuring the Quality of Life Across Countries: A Multidimensional Analysis*. Westview Press.
109. Soteriou, A. and S. A. Zenios (1999), "Operations, Quality, and Profitability in the Provision of Banking Services," *Management Science*, Vol. 45, No. 9, 1221-1238.
110. Stein, Charles (1993), "Grading The Docs," *Boston Globe*, August 15, 77-78.
111. Thompson, R.G., L.N. Langemeier, C.T. Lee, E. Lee and R.M. Thrall (1990), "The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas Farming," *Journal of Econometrics*, Vol. 46, 93-108.
112. Thrall, R.M. (1996), "Duality, Classification and Slacks in DEA," *Annals of Operations Research,* Vol. 66, 109-138.
113. Troutt, M.D., A. Rai, and A. Zhang (1996), "The Potential Use of DEA for Credit Applicant Acceptance System," *Computers and Operations Research*, 23, 405-408.
114. Wennberg, J.E. and A. Gittlesohn (1982), "Variations In Medical Care Among Small Areas," *Scientific American,* 246 (4), 120-134.
115. Yue,P. (1992), "Data Envelopment Analysis and Commercial Bank Performance: A Primer with Applications to Missouri Banks," *Federal Reserve Bank of St. Louis Review*, Vol. 74, No. 1, 31-45.
116. Zhu, J. (1996), "Data Envelopment Analysis with Preference Structure," *Journal of Operational Research Society,* Vol. 47, 136-150.
117. Zhu, J., (2003), *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver*. Kluwer Academic Publishers, Boston.

118. Zhu, J., and Z. Shen (1995), "A Discussion of Testing DMUs' Returns to Scale," *European Journal of Operational Research*, Vol. 81 (1995), 590-596.

INDEX

E

F

G

H

I

L

M

N

O

P

ABOUT THE AUTHORS

h.sherman@neu.edu

College of Business Adminstration
Northeastern University
Boston, MA 02115

H. David Sherman is Professor of Accounting at the Northeastern University, Boston, MA and Adjunct Professor at the Tufts University School of Medicine. He was previously on the faculty of the MIT Sloan School of Management, and was visiting faculty at INSEAD in Fontainebleau, France (1999-2002). In 2004-2005, David was the Academic Fellow at the Division of Corporate Finance of the Securities & Exchange Commission.

Professor Sherman received his MBA and Doctorate degrees from Harvard Business School and his AB in economics from Brandeis University. He is a Certified Public Accountant (CPA) and practiced with Coopers & Lybrand (now PWC).

Professor Sherman began working with DEA as a Doctoral Student. His dissertation explored applications of DEA to hospitals. Professor W.W. Cooper served as his dissertation chair. In addition to health care productivity research, Prof. Sherman has managed DEA applications in banks (including three of the 50 largest US banks), brokerage firms (including the largest discount brokerage firm) and government organizations (including Government of Canada).

Professor Sherman's other research has long focused on financial reporting, performance measurement/management and financial literacy issues facing corporate management and Boards of Directors in global businesses. He served as a director and manager of several businesses, and has consulted on financial performance issues with major corporations including BankBoston; KLA; A.T.Kearney (BankAmerica); U.S. Bancorp; Fidelity Brokerage Services (FMR Corp.); as well as the U.S. Department of Defense and the Financial Accounting Standards Board (FASB). He authored (with D. Young and H. Collingwood): *Profits You Can Trust: Spotting and Surviving Accounting Landmines* (2003-Prentice Hall-Financial Times and 2004 Chinese edition).

Professor Sherman's research has been published in academic and management journals including: *Harvard Business Review* (three articles), *Sloan Management Review, Accounting Review, Interfaces, Management Accounting, Journal of Banking and Finance, American Banker, Medical Care, Auditing. Corporate Board Member, D&O Advisor, Bank Accounting and Finance, Annals of Operations Research and European Journal of Operations Research.* He has been a speaker at business and academic meetings in the U.S., Canada, Europe, Asia and India.

jzhu@wpi.edu

Department of Management
Worcester Polytechnic Institute
Worcester, MA 01609

Joe Zhu is Associated Professor of Operations, Department of Management at Worcester Polytechnic Institute, Worcester, MA. His research interests include issues of performance evaluation and benchmarking, supply chain design and efficiency, and Data Envelopment Analysis. He has published over 50 articles in journals such as *Management Science, Operations Research, IIE Transactions, Annals of Operations Research, Journal of Operational Research Society, European Journal of Operational Research, Information Technology and Management Journal, Computer and Operations*

Research, *OMEGA, Socio-Economic Planning Sciences*, *Journal of Productivity Analysis*, *INFOR*, *Journal of Alternative Investment* and others. He is the author of *Quantitative Models for Evaluating Business Operations: Data Envelopment Analysis with Spreadsheets* (Kluwer Academic Publishers, 2003). He developed the *DEAFrontier* software which is a DEA add-in for Microsoft Excel. Professor Zhu has also co-authored two books on modeling performance measurement and evaluating hedge funds. He is a co-editor of the *Handbook on Data Envelopment Analysis* (Springer, 2004). He is an Associate Editor of OMEGA and The Asia-Pacific Journal of Operational Research. He is also a member of Computers & Operations Research Editorial Board. For more information on his research, please visit www.deafrontier.com.

ABOUT THE CD-ROM

This CD contains the *DEAFrontier* software and the example files. For software installation, please see chapter 4. For more information on the *DEAFrontier* software, please visit www.deafrontier.com.

LICENSING AND WARRANTY AGREEMENT

READ THIS: Do not install or use the CD-ROM until you have read and agreed to this agreement. By opening the accompanying software packet, you acknowledge that you have read and accepted the following terms and conditions. If you do not agree and do not want to be bounded by such terms and conditions, do not install or use the software "*DEAFrontier*".

License: The copyright to the software in the CD-ROM (the "DEAFrontier") is owned by Joe Zhu. The "*DEAFrontier*" is protected by the United States copyright laws and international treaty provisions. No part of the *DEAFrontier* may be reproduced, stored in a retrieval systems, distributed (including but not limited to over the Internet), modified, decompiled, reverse engineered, reconfigured, transmitted, or transcribed, in any form or by any means without the permission of the author. The *DEAFrontier* may not, under any circumstances, be reproduced for sale. **This license allows you to use the *DEAFrontier* for educational and research purposes only, not for commercial purposes, including consulting services.** You may only (i) make one copy of the DEAFrontier for backup or archival purposes, or (ii) transfer the DEAFrontier to a single hard disk, provided that you keep the original for backup or archival purposes.

You may not (i) rent or lease the *DEAFrontier*, (ii) copy or reproduce the *DEAFrontier* through a LAN or other network system or through any computer subscriber system, or (iii) modify, adapt, or create derivative works based upon the *DEAFrontier*. You may be held legally responsible for any copying or copyright infringement which is caused by your failure to abide by the above terms and conditions.

Limited Warranty: The warranty for the media on which the DEAFrontier is provided is for thirty (30) days from the original purchase and valid only if the packaging for the DEAFrontier was purchased unopened. If the Publisher receives notification within the warranty period of defects in materials or workmanship, the Publisher will replace the defective software media. The Publisher and the author provide no other warranties, expressed or implied, including without limitation implied warranties of merchantability and fitness for a particular purpose, and shall not be liable for any damages, including direct, special, indirect, incidental, consequential, or otherwise. The Publisher does not warrant that the functions contained in the DEAFrontier will meet your requirements or that the operation of the DEAFrontier will be error-free.

SOFTWARE INSTALLATION

Please see chapter 4 for detailed instruction.

To install the software the CD-ROM using Windows, you may follow these steps:

Step 1. Insert the CD-ROM into your computer's CD-ROM drive.

Step 2. Launch Windows Explore.

Step 3. Click Browse to browse the CD and find the file "*DEAFrontier.xla*". Copy this file to your hard drive (please see footnote 2 in chapter 4)

DEAFrontier does not set any limit on the number of units, inputs or outputs. However, please check www.solver.com for problem sizes that various versions of Excel Solver can handle (see Table 4-1).

To run *DEAFrontier*, the Excel Solver must first be installed, and the Solver parameter dialog box must be displayed at least once in the Excel session. Otherwise, an error may occur when you run the software. *Please make sure that the Excel Solver works properly. One can use the file "solvertest.xls" to test whether the Excel Solver works. This test file is also available at www.deafrontier.com/solvertest.xls.*