

## 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

Car	# miles traveled	MPG	Actual Inputs			Actual Output
			Cost per gallon	Total Cost	Cost per mile	# of miles
<i>John</i>						
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	
<i>Mary</i>						
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 Unit	Service Output		Input Used	
	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.

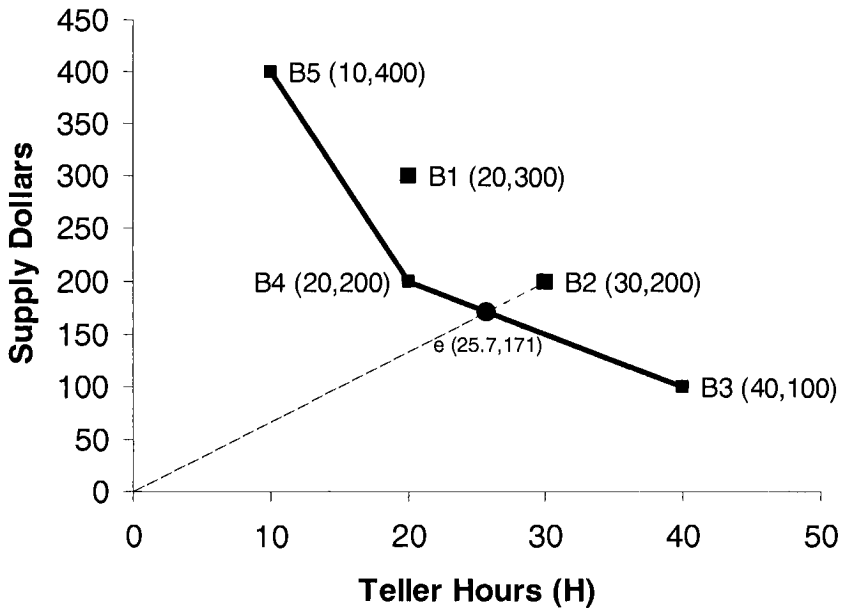


Figure 2-1. Graphic Representation of the Five Bank Branches

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 ( $\theta < 1$ )<sup>1</sup>. DEA will provide 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

<sup>1</sup> The symbol  $\theta$  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 ( $\theta$ )	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 ( $\theta$ ). 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 B5 have 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 ( $0.857 \times 200$ ) and use of 171 supply dollars ( $0.857 \times 200$ ). 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 than B2. These ERS weights are generally referred to as Lambda -  $\lambda$  - 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)	$\begin{bmatrix} 1,000 \end{bmatrix}$	$\begin{bmatrix} 1,000 \end{bmatrix}$	$\begin{bmatrix} 1,000 \end{bmatrix}$
<u>Inputs</u>	$(0.2857) \times$	$+ (0.7143) \times$	$=$
Teller Hours (H)	$\begin{bmatrix} 40 \end{bmatrix}$	$\begin{bmatrix} 20 \end{bmatrix}$	$\begin{bmatrix} 25.7 \end{bmatrix}$
Supply \$(S)	$\begin{bmatrix} 100 \end{bmatrix}$	$\begin{bmatrix} 200 \end{bmatrix}$	$\begin{bmatrix} 171 \end{bmatrix}$

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

	<u>Column 1</u> Composite Outputs & Inputs (from above)	<u>Column 2</u> Branch B2 Actual Outputs & Inputs	Column 2 – Column 1
(T)	1,000	1,000	0
(H)	25.7	30	4.3
(S)	171	200	29

$\left. \begin{matrix} \text{Excess Inputs} \\ \text{Used by} \\ \text{Branch B2} \end{matrix} \right\}$

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

SU $j$  = service unit number  $j$

$\theta$  = 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  $\theta$  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

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

(Maximize the efficiency rating  $\theta$  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:

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

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

...

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

...

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

$$u_1, \dots, u_s > 0 \text{ and } v_1, \dots, v_m \geq 0$$

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:

$$\begin{array}{ll} \text{B1} & \frac{u_1(1000)}{v_1(20) + v_2(300)} \leq 1 \\ \text{B2} & \frac{u_1(1000)}{v_1(30) + v_2(200)} \leq 1 \\ \text{B3} & \frac{u_1(1000)}{v_1(40) + v_2(100)} \leq 1 \\ \text{B4} & \frac{u_1(1000)}{v_1(20) + v_2(200)} \leq 1 \\ \text{B5} & \frac{u_1(1000)}{v_1(10) + v_2(400)} \leq 1 \end{array}$$

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,  $v_1 = 0.01429$  for teller hours (H) and  $v_2 = 0.00286$  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  $\times$  1.43%=14.3%) or by decreasing S by 50 units (50 $\times$ 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<sup>2</sup>:

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

Subject to the constraints that<sup>3</sup>

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

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

<sup>2</sup> This expression maximizes the numerator for the unit being evaluated, trying to assign it the highest possible productivity rating.

<sup>3</sup> 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:

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

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

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

subject to

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, j = 1, \dots, n \quad (2.1)$$

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

$$u_r, v_i \geq 0$$

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

$$\sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, 2, \dots, m; \quad (a) \quad (2.2)$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad r = 1, 2, \dots, s; \quad (b)$$

$$\lambda_j \geq 0 \quad j = 1, 2, \dots, n. \quad (c)$$

The dual is seeking the efficiency rating, minimize  $\theta$ , 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  $\lambda$  (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  $\lambda$  value for B3 was 0.2857 and for B4 was 0.7143. The efficiency rating  $\theta$  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  $\theta$  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  $\theta$  for the service unit being evaluated and it cannot find  $\lambda$  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 of 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 ( $\lambda_j$ ); (2) cell for the objective function (efficiency,  $\theta$ ); (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  $y_{ro}$ )

1	A	B	C	D	E	F	G	H
2	Service Unit	Service Output		Input Used				
3	B1	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)		$\lambda$	changing cell
4	B2	1,000		20	300		1	changing cell
5	B3	1,000		30	200		0	changing cell
6	B4	1,000		40	100		0	changing cell
7	B5	1,000		20	200		0	changing cell
8				10	400			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  $\lambda_j$  ( $j = 1, 2, \dots, 7$ ). Cell F12 represents the efficiency score  $\theta$ , 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  $\lambda_j$  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<sup>4</sup>.

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  $\lambda_j$  s 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

<sup>4</sup> 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  $\lambda$  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  $\theta$  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  $\theta$  value with the “dual” will equal the maximum  $\theta$  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.

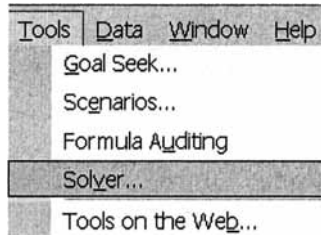


Figure 2-3. Display Solver Parameters Dialog Box

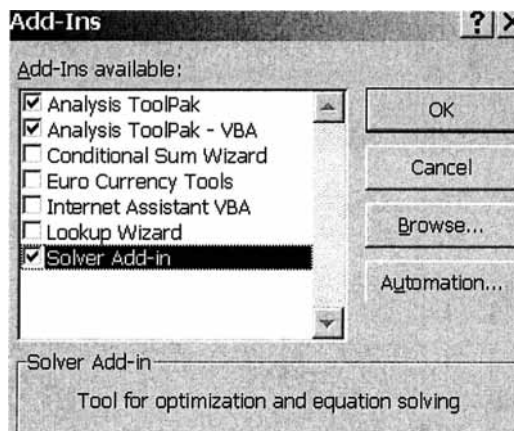


Figure 2-4. Solver Add-In

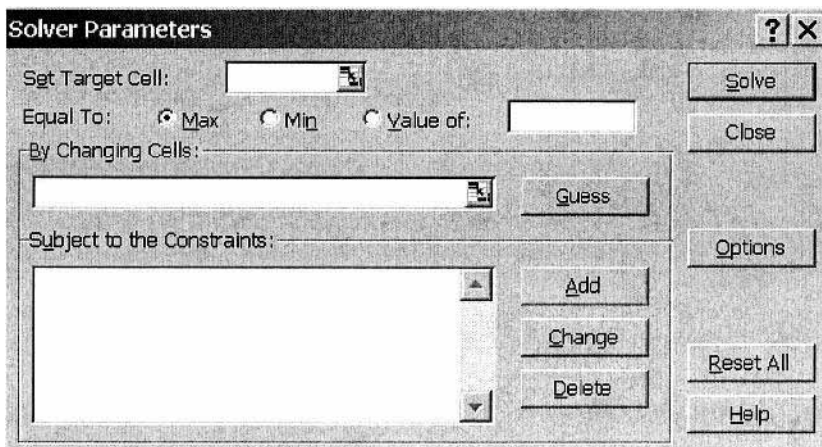


Figure 2-5. Solver Parameters Dialog Box

### 2.4.3.2 Specifying the Target Cell

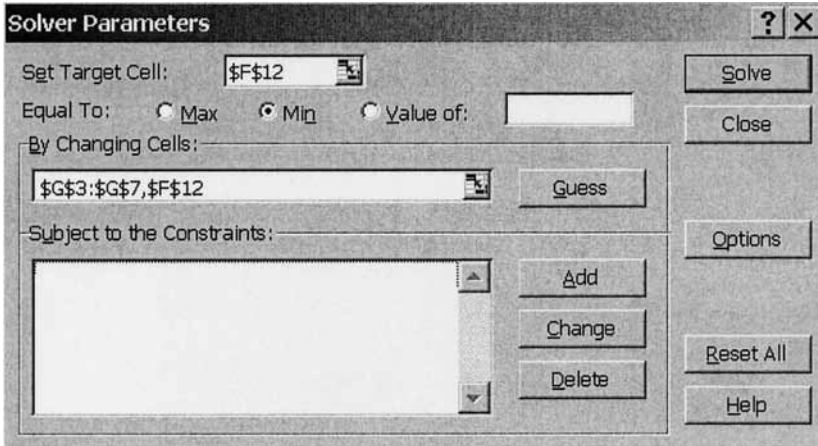


Figure 2-6. Specifying Target Cell and Changing Cells

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  $\lambda_j$  ( $j = 1, 2, \dots, 7$ ) and  $\theta$ , and should be cells G3:G7 and F12, respectively (see Figure 2-6).

### 2.4.3.4 Adding Constraints

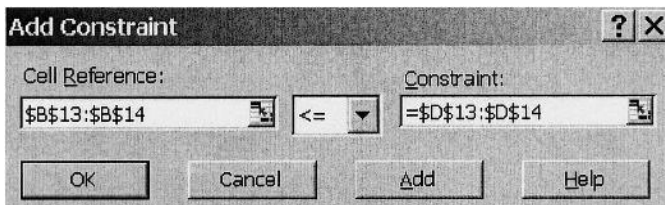


Figure 2-7. Adding Constraints

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  $\lambda_j$  and  $\theta$  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.*

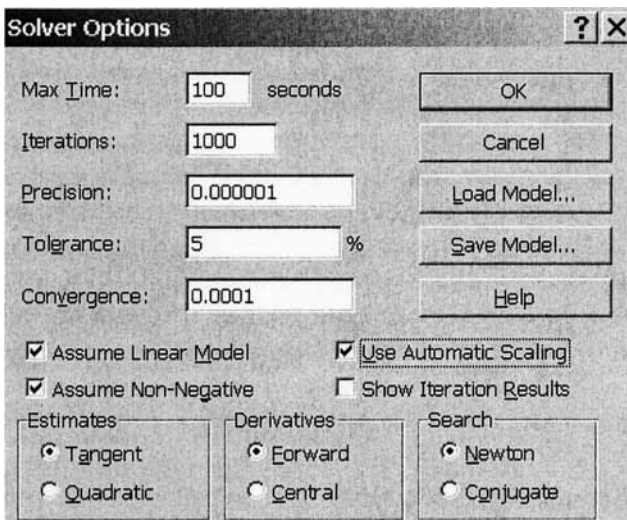


Figure 2-8. Non-Negative and Linear Model

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.

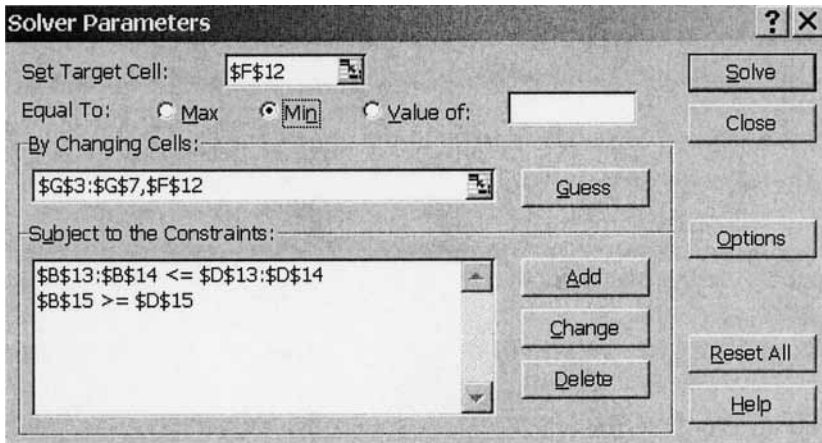


Figure 2-9. Solver Parameters for DEA Model (2.2)

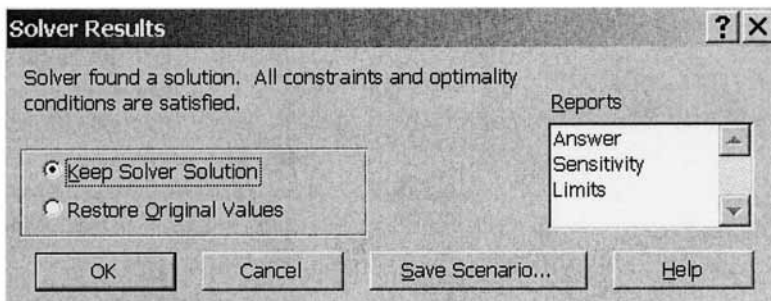


Figure 2-10. Solver Results Dialog Box

	A	B	C	D	E	F	G
1		Service Output		Input Used			
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars (S)		$\lambda$
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  $\lambda$  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.xls** (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.)

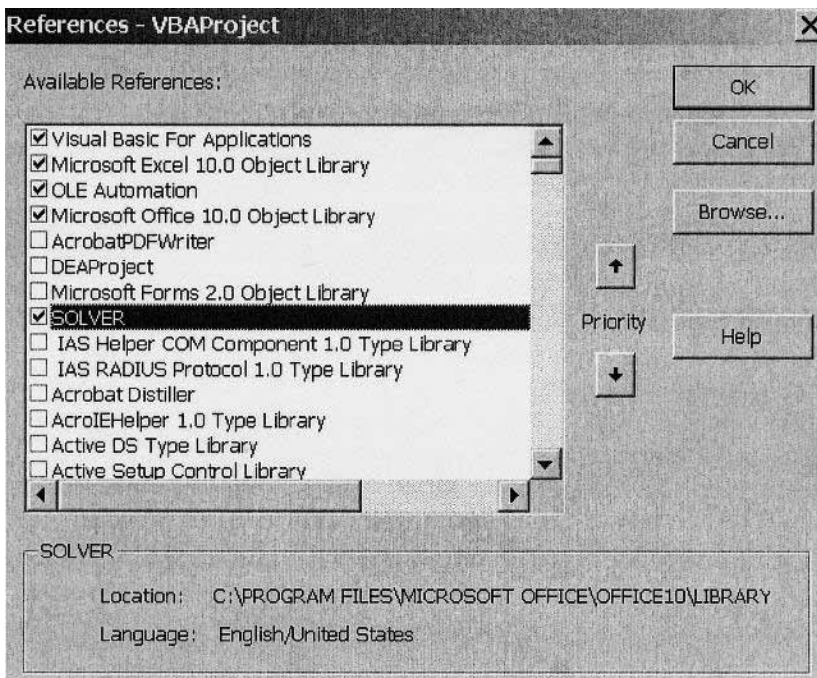


Figure 2-12. Adding Reference to Solver Add-In

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.)



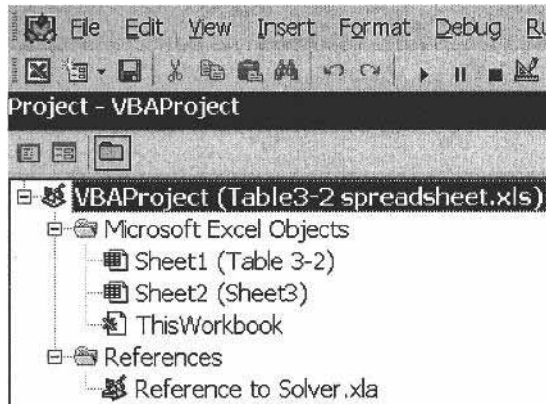


Figure 2-13. Reference to Solver Add-In in VBA Project

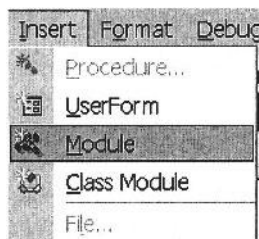


Figure 2-14. Insert a Module

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).

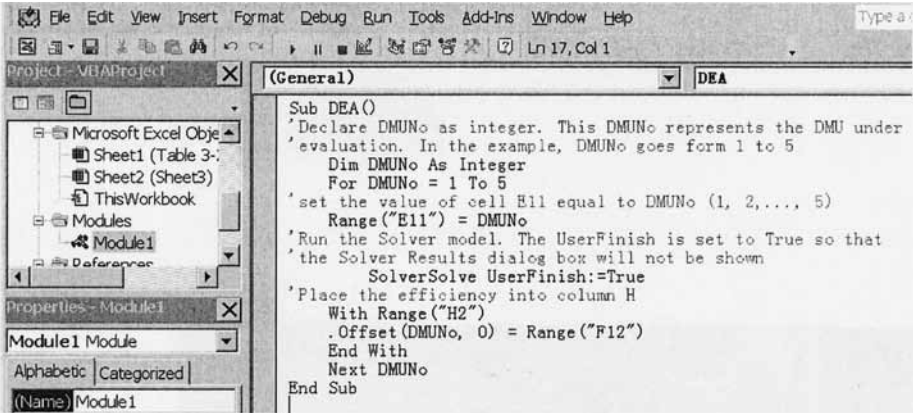


Figure 2-15. VBA Code for Envelopment Model

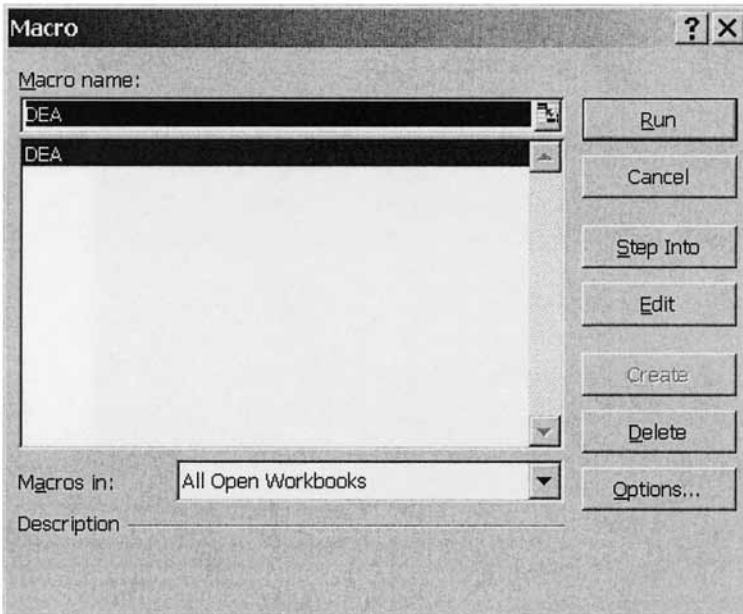


Figure 2-16. Run "DEA" Macro

	A	B	C	D	E	F	G	H
1		Service Output		Input Used				
2	Service Unit	Transactions Processed (T)		Teller Hours (H)	Supply Dollars		$\lambda$	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  $\lambda_j$  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
```

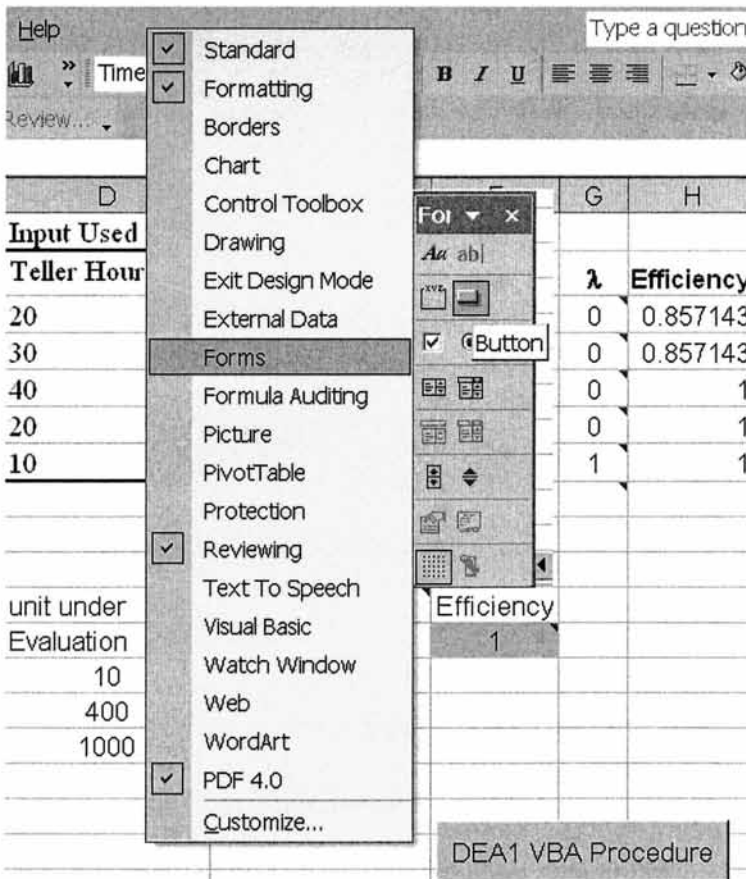


Figure 2-18. Adding a Button with Macro

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  $\lambda_j$  ( $j = 1, 2, \dots, 5$ ) (see file “Table2-2 spreadsheet.xls” in the CD). This will generate the  $\lambda_j$  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,  $\lambda_j$  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  $\lambda$  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.

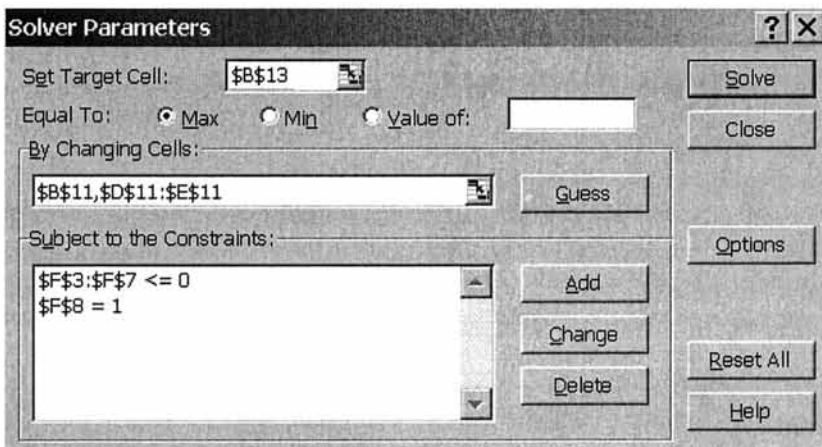


Figure 2-20. Solver Parameters for Multiplier Model

	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 resolve the problem. Figure 2-21 shows the results for branch B2. Note that we have  $u = 0.0857 (\times 10^{-2})$  in cell B11,  $v_1 = 1.429 (\times 10^{-2})$  in cell D11 and  $v_2 = 0.286 (\times 10^{-2})$  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 *DEA Frontier*. The *DEA Frontier* does not set limit on the number of units, inputs or outputs. With the capacity of Excel Solver, the *DEA Frontier* can deal with large sized performance evaluation tasks.<sup>5</sup>

<sup>5</sup> 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.



<http://www.springer.com/978-0-387-33211-6>

Service Productivity Management  
Improving Service Performance using Data Envelopment  
Analysis (DEA)

Sherman, H.D.; Zhu, J.

2006, XXII, 328 p. With CD-ROM., Hardcover

ISBN: 978-0-387-33211-6