Modeling Undesirable Outputs in Data Envelopment Analysis: Various Approaches

Kalyan Sunder Pasupathy

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Konstantinos P. Triantis, Ph.D., Chair
Barbara J. Hoopes, Ph.D.
C. Patrick Koelling, Ph.D.

February 19, 2002
Falls Church, Virginia

Key Words: Data Envelopment Analysis, Performance Measurement, Undesirable Outputs, Technological Dependence, Goal Programming

Modeling Undesirable Outputs in Data Envelopment Analysis: Various Approaches

Kalyan Sunder Pasupathy

(ABSTRACT)

The general practice in performance and production efficiency measurement has been to ignore additional products of most transformation processes that can be classified as "undesirable outputs" — which are a subset of the output set. Without the inclusion of these factors, the efficiency evaluation becomes a purely technical measure of the system alone, and does not account for the interaction of the system with the surrounding environment and the impact of policy decisions on the system. In addition, there are also technological dependencies arising due to the relationships between the desirable and the undesirable outputs. One of the analytical tools normally used in efficiency evaluation is Data Envelopment Analysis, DEA.

In the course of addressing these problems, a decision-maker encounters multiple and contradictory objectives with respect to the output set. This motivates the exploration of new arenas of measurement of efficiency to facilitate policy decisions and address technological relationships. This research presents five modifications of the traditional DEA technique to give a more realistic and comprehensive score of production efficiency considering both, desirable and undesirable outputs. The models address the following problems: (i) technological dependency between desirable and undesirable outputs; (ii) decision-maker's preferences over inputs, desirable outputs and undesirable output performance and finally (iii) conflicting production objectives through a formulation that uses Goal Programming in conjunction with DEA, a concept known as GoDEA.

To

My Mom and Dad

ACKNOWLEDGEMENTS

I wish to dedicate this Thesis to my Mom Mrs. Lalitha Pasupathy and my Dad Dr. N. K. Pasupathy, who were there for me in all trials and tribulations all through my life. It would not be appropriate not to dedicate this work also to my Maternal Grandparents Mrs. Padma Balakrishnan and Mr. K. S. Balakrishnan, who were like my second parents and with whom I spent a great part of my childhood. Credits for the fact that I'm here and am able to accomplish a task go to these four important people in my life. I also thank my younger brother Arrvind Sunder for being more of a friend and a companion in my life.

My Advisor Dr. Triantis, the "Master", deserves all tribute for opening my eyes and for stimulating me to look at things from a broader perspective. He was (and still is) a very close friend in the entire journey of research at whose door I could go knocking anytime.

Thanks to my other Committee members Dr. Hoopes and Dr. Koelling, for helping me improve display my work. I should also thank Mrs. Lovedia Cole who greets me happily and is always ready to help even when I went to her in the "eleventh" hour especially for a student like me in an extended campus. Many thanks to all of you friends for lightening the burden and cheering me all the way up the road.

Thank you God! for giving me the will power and the fortitude to leap all hurdles.

Am I forgetting somebody here? Of course. My very best and special friend whom I have known for just more than a year- Alex. She has been my philosopher, guide and mentor not only for this Thesis but also for life. A person whom I could go to for moral and emotional support. How can I put out of my mind the numerous intense debates and the hours of endless discussions with her in the GA's Office? Without her endurance and perseverance this whole process could not at all have been viable.

Table of Contents

List of Figures

List of Tables

CHAPTER 1. Introduction

1.1. OBJECTIVES

The main objectives of this research are:

1. To explore new model formulations for including undesirable outputs in Data Envelopment Analysis. These formulations will attempt to

- Find a way of maximizing desirable outputs while minimizing undesirable outputs in efficiency computations.
- Determine performance measures for the firms in the presence of predetermined technological dependence between desirable and undesirable outputs.
- Develop a new approach that determines the linear dependence of the desirable and the undesirable outputs and computes efficiency score for the firms based on their performance with respect to the two types of outputs.
- Develop a modified model with a preferential weight structure to set targets for the firms in the presence of the desirable and the undesirable outputs.
- Finally to explore a Goal Programming approach that addresses the issue of multi-objective problems relating to inputs, desirable outputs and undesirable outputs. This would be the first time that this technique is used to evaluate production units with conflicting production goals. This is also a suitable approach for setting production targets.

2. To validate and verify the suitability of the proposed models through:

- Consultation with decision makers to validate the results
- Comparison with previous results from existing models
- Peer review through the presentation of the model in conferences where DEA experts participate (INFORMS 2001).

1.2. MOTIVATION

There have been hundreds of papers addressing the issues related to performance measurement and production efficiency. Most of these papers consider solely inputs or resources used by a firm and the desirable outputs or operational products that are the result of input utilization. Other production variables are not included in the traditional model formulation. These include environmental variables such as pollution, various types of undesirable outputs such as scrap, rework, and other qualitative outputs such as service characteristics that lead to dissatisfied customers, to name a few. Without the inclusion of these factors, the evaluation of technical performance ignores real world considerations.

The undesirable outputs are an anomaly in the whole set of outputs, which makes them interesting and worthwhile to analyze. Notwithstanding, the fact that the nature of these types of outputs is different from that of the desirable outputs, they demand a different set of assumptions related to the production possibility set and the modeling of the production process.

Furthermore, the problem of incorporating undesirable outputs into efficiency measurement requires to reward units that produce more quantity of desirable outputs and less quantity of undesirable outputs. This multi-objective nature of the problem requires the utilization of appropriate techniques to handle these differences. Further, the differences in behavior and target goals between the two types of outputs need to be analyzed and understood well before they can be expressed mathematical terms. Including the undesirable outputs necessitates research work on how to account for them along with the desirable outputs in the production formulation.

The effect of managerial or policy decisions on the transformation process can be expressed in different forms according to the type of problem and system being analyzed. In the case of emissions or pollution, for example, regulatory standards are placed on the amount of undesirable outputs obtained on the output side of the production process. In the case of a manufacturing process, technological constraints can impose a certain minimum amount of scrap that is generated independently of the amount of good output produced. It can also be the case that a company's policy could limit the amount of undesirable (bad) output produced even to the detriment of desirable (good output) maximization goals. Addressing this problem of conflicting production goals is an additional step in the direction of taking the modeling world closer to the real world.

In addition, the classical DEA formulation faintly recognizes the effect of the production process on the environment or the impact that the surroundings have on the performance of the system. In the research work being addressed in this document, ways are analyzed on how to approach the issues described above.

1.3. PRIOR APPROACHES

There have been different directions in which the problem of modeling undesirable outputs for efficiency evaluation has been explored.

Some authors have approached the problem through transformations performed on the undesirable output data so that these could be considered just as desirable outputs. Then, the transformed undesirable output data is considered to have the same characteristics as the desirable output data. Scheel (2001) discussed various transformations of the undesirable output data. These transformations will be discussed later in Chapter 2 of this document. Some of the drawbacks of these approaches are that because of the transformation done to the data set, we might run into problems associated with convexity and non-linearity assumptions in DEA. Also, in these cases the framework assumes that the transformed data has its own meaning, like the undesirable output "mortality rate" and its additive inverse "survival rate". In many other real life applications, the transformation of the data may not make sense.

Other ways of including undesirable outputs is to modify the underlying production assumptions. Thus at the modeling stage, the two different types of outputs are incorporated differently in the formulation. In this approach the data do not need to be transformed.

Among these approaches, one method developed by Fare et al. (1989) defines a hyperbolic path to determine a performance measure. This type of measure, which is discussed in detail in Chapter 2, attempts to increase the desirable outputs and decrease the undesirable outputs in the same model. This approach involves defining a non-linear projection due to the hyperbolic path defined by the model. Fare et al. (1989) uses a linear approximation to address the non-linear projection. However, there are problems associated with this linear approximation which is discussed in section 2.7.3.

Zofio and Prieto (2001) came up with a modified model with a transformation of the same expression to overcome the problems with the linear approximation. So far, we have been limiting our discussions to single stage approaches, where both the increase of the desirable outputs and the decrease of the undesirable outputs are formulated in the same model in one single step. The Index number approach, which is discussed next, is a two-stage model.

Fare et al. (2000) formulated an Index number approach that evaluates the performance of the firms using two separate models, one in each stage. The first stage computes an index based on the increase of the desirable outputs and the second stage based on the decrease of the undesirable outputs. Then an overall environmental index is computed using the indices from the two stages.

So far, we have briefly mentioned the various approaches in the literature for modeling undesirable outputs. In the next section, we will present the motivation that are the basis for this research and support the proposed methodology for the development of new DEA formulations.

1.4. PROPOSED METHODOLOGY

As it can be seen, the problem at hand includes more than one objective-maximization of the desirable outputs, minimization of the undesirable outputs. Each of these objectives may have relative importance. Further, within the sets of each of the two types of outputs, different variables may have varying importance. This motivates one to formulate a model to include the relative importance of the objectives as well as the relative importance of the variables.

Thus, a model with a preferential weight based structure for the desirable outputs and the undesirable outputs is developed and will be utilized to measure the performance of the DMUs and estimate targets that account for the level of desirable as well as undesirable outputs in the model. The concept behind this model is based on the work by Thanassoulis and Dyson (1992).

Further, along with the two types of outputs, there are also inputs in the transformation process that need attention. A decision-maker may encounter the problem to minimize the inputs along with the two objectives above. Again, each of these inputs may have varying degrees of importance. Considering the inputs as well, one is looking at a multi-objective problem that could be solved by using Goal Programming in DEA. This technique —called GoDEA—was presented by several authors (Charnes et al. (1988), Thanassoulis and Dyson (1992), Athanassopoulos (1995), Sheth (1999), Hoopes, Triantis and Partangel (2000)) and was developed to combine conflicting objectives of resource allocation and also to encapsulate the position of different levels of management in the generation of planning scenarios. The nature of multi-objective programming methods has been advocated to solve problems where different goals are conflicting (such as desirable output maximization and undesirable output minimization and input minimization at the same time for the same firm).

Another issue that this research will attempt to address is to consider the technological relationship existing between good and bad outputs. This relationship exists in majority of manufacturing processes as well as in many other processes in other realms. For example, the scrap resulting from molding or plastic injection, or chemical waste products associated with process industries are representative of not only of the existence of undesirable outputs but their technological relationship to the good outputs is known or can be calculated.

The existing DEA approaches that treat undesirable outputs do not address how these type of relationships affect efficiency performance. This is another objective of this research work- to account for the technological dependence between the desirable and the undesirable outputs. For example, this amounts to including an additional constraint in the model developed by Zofio and Prieto (2001) and Fare et al. (2000).

However in many instances, the technological dependence is neither known a priori nor can be calculated from the process. In such cases, an approach needs to be developed that can help the decision-maker determine this dependence. Knowing the relationship would help him/her define performance targets for the firms and assess the efficiency performance of the firms based on this technological dependence. To address this problem, we propose a novel approach that determines the technological relationship between undesirable and desirable outputs and the resulting efficiency performance based on this relationship.

In summary, there are three innovations to the traditional DEA formulations that are being proposed:

1. A model with a preferential weight based structure for the desirable outputs and the undesirable outputs;
2. A GoDEA approach to consider multi-objectives for inputs, desirable outputs and undesirable outputs.
3. A modification to existing formulations that considers the technological relationship existing between good and bad outputs.
4. An approach to help the decision-maker to determine this technological dependence between undesirable and desirable outputs when it is unknown, and the resulting efficiency performance based on this relationship.

Results from all four approaches will be compared.

1.5. ORGANIZATION OF THIS DOCUMENT

Chapter 2 starts with an explanation of the fundamental measures in efficiency measurement. It then presents the theoretical background on returns to scale and on the disposability of inputs and outputs in production. Then it presents Data Envelopment Analysis and explains the CCR and BCC models. Various non-radial measures are also discussed. Research related to the treatment of undesirable outputs is reviewed. The various types of measures of efficiency when undesirable outputs are included are also discussed. Finally, the GoDEA approach is presented.

Chapter 3 presents the methodology proposed in this research by describing new potential formulations for the treatment of undesirable outputs. It also provides a brief overview of the data that will be used in illustrating the application of these formulations.

Chapter 4 illustrates the four formulations developed in Chapter 3 for the treatment of undesirable outputs. We do this by applying them to three data sets, two from the DEA literature and another one collected with the purpose of this research. The results obtained from other models will be used to compare the performance of the formulations developed here.

Chapter 5 presents the conclusion of this research effort and make recommendations for future research. The first section summarizes the research of this thesis. The second section describes the major contribution of this research and the third section outlines some recommendations for future research.

Finally, the Appendix has the actual data sets, used in Chapter 4 for the application of the models introduced in Chapter 3. The Appendix also displays the programming code used to run the models in MS Excel Solver and Visual BASIC for Excel.

CHAPTER 2. Literature Review

This chapter contains a summary of all the literature that is relevant to the topics covered in the remainder of the thesis. The topics covered are efficiency measurement, technical efficiency, relative efficiency measurement, Data Envelopment Analysis (DEA), the inclusion of undesirable outputs in efficiency measurement and finally a section in goal programming and DEA.

2.1. THE TRANSFORMATION PROCESS

The whole idea of efficiency measurement relies on production theory, which sees a firm as a production system where inputs are the resources that are utilized by the firm or the organization and are transformed into desirable outputs. This transformation process is depicted in Figure 2.1.

Figure 2. 1 The Transformation Process

2.2. EFFICIENCY MEASUREMENT

The traditional measure of efficiency determines a score based on the ratio of the output that was obtained from the process and the inputs or resources that were used by the process. This traditional efficiency score was thus given as in equation 2.1 below.

$$\text{Efficiency} = \frac{\text{Output}}{\text{Input}} \qquad \textbf{(2.1)}$$

This measure of efficiency had its own drawbacks, some of which are described below.

(i) Inability of the model to incorporate multiple inputs and outputs.

(ii) Real life scenarios that incorporate other process dimensions such as quality and outcomes cannot be easily incorporated in one single equation.

(iii) Environmental factors that affect the process under study cannot be easily modeled.

(iv) As a continuation of (i) above, in the presence of multiple inputs and outputs, varying units of the variables cannot be handled.

2.3. TECHNICAL EFFICIENCY

Considering the above-mentioned drawbacks, Farrell (1957) introduced a new measure of efficiency to take into account all inputs and outputs. This measure was defined in such a way so as to overcome the drawbacks mentioned above, and to know how far "a given industry can be expected to increase its output by simply increasing its efficiency, without absorbing further resources." p. 260

This new measure of efficiency, known as technical efficiency, is a score determined for each firm. The firm is analyzed within a group of comparable firms and is evaluated by comparing it with some ideally performing firm. This ideally performing firm is found by one of the following means.

(i) Theoretical - the entire process is represented as a theoretical or “ideal” production function where the outputs produced by the process are represented as a function of the inputs. This function provides the ideally performing firm and the expected performance from the comparison.

(ii) Empirical - unlike the theoretical approach that is impossible to operationally achieve, the performance of the firm is determined by comparing it to a relative production combination that is achievable in practice.

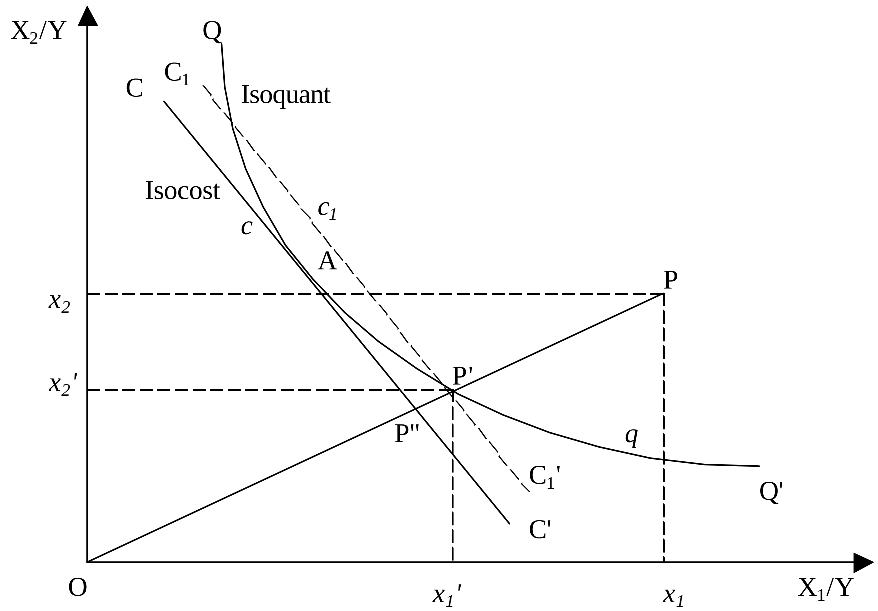


Figure 2. 2 Isoquant: Input-Orientation

2.3.1. Input-Oriented Measures

Let us consider a production process with two inputs X_1 & X_2 and one output Y. Figure 2.2 graphically depicts this production function. QQ' is termed as an **isoquant** and represents the *efficient production function*. An isoquant is the locus of all possible combinations of the inputs that produce the same amount of output. It is convex to the origin showing the lower bounds on the inputs. Before the actual comparison is done, the output Y is held at a fixed quantity of value *q*. In the above Figure 2, the quantity *q* is assumed to be one. All the firms that are to be compared and normalized so that they produce the same amount of output quantity *q* using varying quantities of inputs X_1 and X_2 based on their production capabilities. Once this is done, all the production firms are plotted in the diagram to obtain a scatter plot. The isoquant is then computed or estimated and is also included in the scatter plot. It represents some function that is obtained from a relative combination of all existing firms. It is defined in such a way that all the firms lie to the northeast of the isoquant. P is one such firm represented in the scatter plot.

Firm P utilizes x1 and x_2 units respectively of inputs X_1 and X_2 to produce an output quantity *q*. For the firm P to perform efficiently, it should use x_1' and x_2' units of inputs respectively to produce the same quantity *q* of the output Y. In other words, the firm P would have been operating efficiently, if it uses x_1 and x_2 units of the inputs to produce *q* units of the output Y. Hence based on the former discussion, where the inputs are reduced proportionally holding the output constant, the **technical efficiency** of the firm P is given as OP'/OP. Typically, this means that the two inputs could be reduced by a proportion equal to OP'/OP. Reducing input X_1 by a proportion OP'/OP from x_1 means to bring it down to x_1' units (by the law of similar triangles). A similar proportionate reduction is done for the input X_2 to reduce it from x_2 to x_2'.

In addition to the technical efficiency, the cost of the inputs should also be considered to determine the overall performance of the firm under investigation. The line CC' is the **isocost** line representing the various combinations of the two inputs that have the same total cost *c*. The slope of the isocost line is determined by the ratio of the costs of the two inputs. Higher cost lines that have the same cost ratio as CC' lie to the northeast of CC' and parallel to it.

Since the isocost line CC' is tangential to the isoquant QQ' at point A, the firm that is determined by the point A would essentially have the best technical and **allocative efficiency**. Allocative efficiency portrays the ability of the firm to use the inputs in optimal proportions so that the resource cost is minimized. Firm P' which is also the projection of firm P onto the isoquant QQ' is also as technically efficient as firm A, but is not as allocatively efficient as A. This is because, the cost of production at point P' would be the cost associated with the isocost line C_1C_1', which is c_1. As per definition, the cost c_1 is higher than the cost *c*. The allocative efficiency of the firms P and P' is the ratio OP'/OP.

By definition (Farrell (1957)), the **total economic efficiency** of the firm P is the ratio OP''/OP. This ratio is defined as follows:

$$OP''/OP = OP''/OP' \times OP'/OP$$

Hence,

Total economic efficiency = Allocative efficiency x Technical efficiency **(2.2)**

For the firm P', the total economic efficiency would be OP''/OP and this would be the same as the allocative efficiency OP''/OP' since the technical efficiency of firm P' is one (1). All the three measures have an upper limit of one and a lower limit of zero.

The above procedure assumes that the production function of the firm is known. However, in most practical scenarios, this is not the case. The production function is either too complicated to be represented or may not be known at all. In these circumstances, Farrell (1957) suggested the use of a non-parametric piece-wise linear convex isoquant such that in either case, no firm lies either to the left or to the bottom of the isoquant. Such a function envelops all the data points as shown below in Figure 2.3.

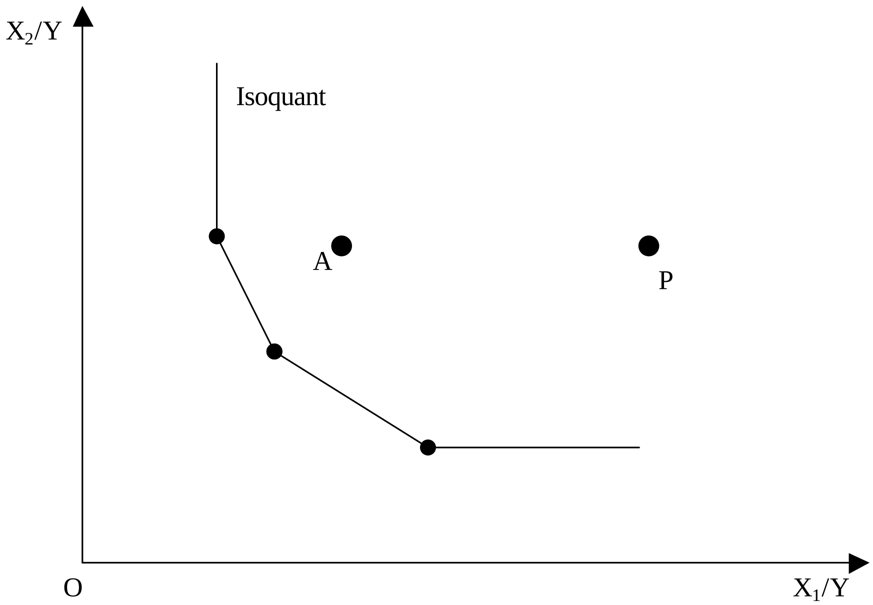


Figure 2. 3 Piecewise Linear Isoquant

2.3.2. Output-Oriented Measures

The previously discussed technical efficiency measure determined the efficiency of a firm based on how much the inputs could be reduced proportionally without a decrease in the output and hence is known as the input-reducing measure of technical efficiency. On the other hand, the output-oriented measure looks at the extent to which the outputs produced can be increased without an increase in the inputs and hence is known as the output-increasing measure of technical efficiency. This can be illustrated using Figure 2.4. Here a single input X is involved in the production of two outputs Y_1 and Y_2. ZZ' is the isoquant that represents a constant quantity of input that is used to produce varying proportions of the two outputs. It is concave to the origin and since it determines the best production possibility, all the firms lie to the left and the bottom of ZZ'. A is one such firm. Point B is the projection of the firm A onto the isoquant ZZ'. Here the distance AB determines the amount of technical inefficiency. Hence the output-oriented technical efficiency measure is given by OA/OB. If the prices of the outputs are also known, then the **isorevenue** line EE' can be drawn and the allocative efficiency can also be determined as OB/OD. Then the overall efficiency would be the product of the two efficiencies and would be the ratio OA/OD.

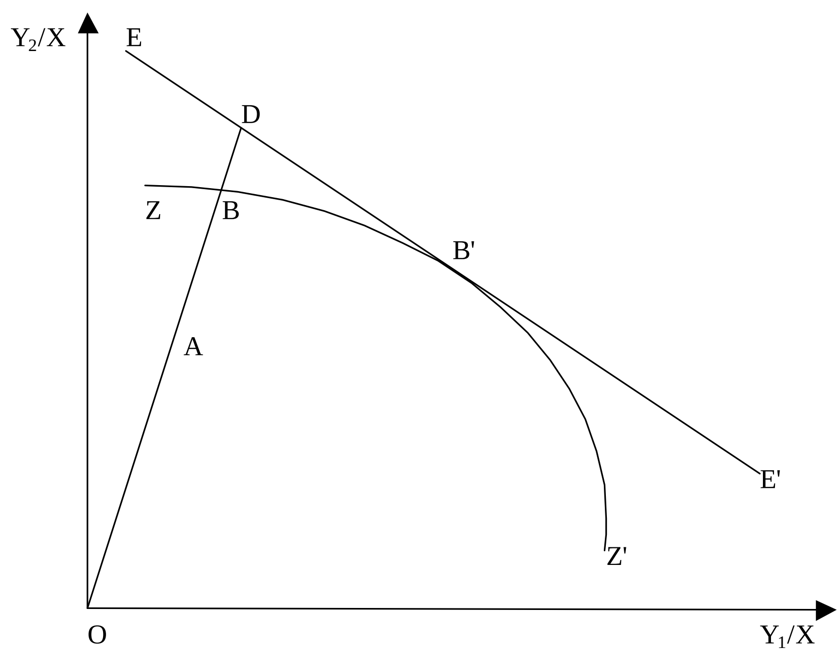


Figure 2. 4 Output-Orientation

2.3.3. Returns to Scale

The above efficiency measures are based on **constant returns to scale** technology (CRS). This implies that the production technology under consideration is such that, an increase in all the inputs by some proportion results in an increase in the outputs by the same proportion. The **non-constant (or variable) returns to scale** result in a non-proportionate change (increase or decrease) in the outputs. The three types of returns to scale and the difference between the input-reducing and the output-increasing measures are illustrated on Figure 2.5.

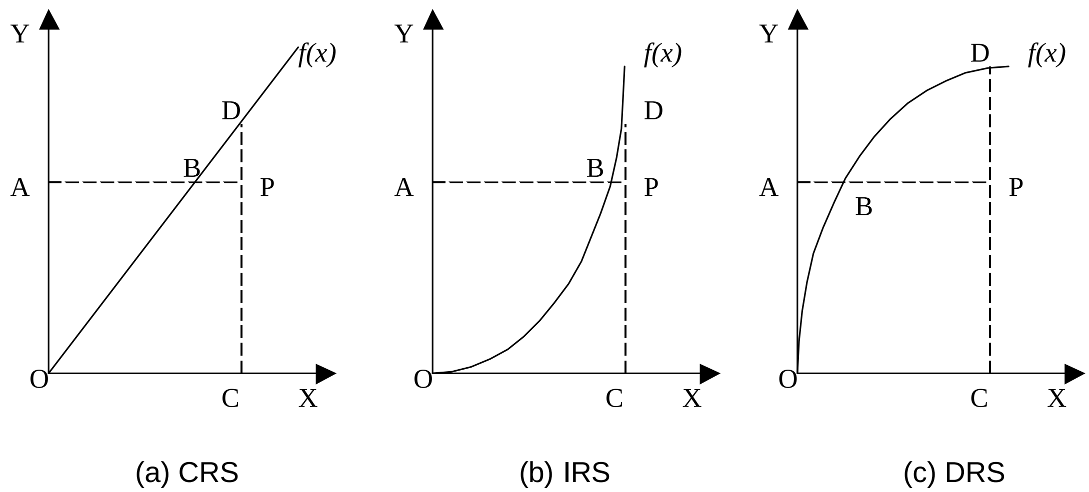


Figure 2. 5 Returns to Scale

In Figures 2.5, a production of a single output from a single input is illustrated graphically. In Figure 2.5 (a), it can be seen that the function *f(x)* is a straight line and has a single slope. Hence for every unit increase in the input that goes into the process, the output produced increases by a constant proportional quantity, hence it represents constant returns to scale (CRS). Figure 2.5 (b) represents a function with an increasing slope. For every unit increase in the input, the output increases by a more than proportional quantity displaying increasing returns to scale (IRS). Finally Figure 2.5 (c) represents decreasing returns to scale (DRS) where the function has a slope that decreases as the input increases. P is a firm that lies below the efficient frontier. In each of the three cases, P could be projected onto the frontier either under an input-reducing consideration or an output-increasing consideration. B and D are projected points on the frontier obtained for comparison. The input-reducing efficiency measure is given by AB/AP and the output-increasing efficiency measure is given by CP/CD.

In the case of constant returns to scale as in Figure 2.5 (a), the triangles ΔOAB and ΔDCO are similar. By the law of similar triangles,

$$AB/OC = OA/CD$$

$$\Rightarrow \quad AB/AP = CP/CD$$

Hence, the input-reducing and the output-increasing measures give the same technical efficiency score under the assumption of constant returns to scale. However, in the case of variable returns to scale, both measures give different efficiency scores. It can also be seen that the increasing returns to scale assumption tends to increase the distances either AB or CD. In Figure 2.5 this results in an input-reducing efficiency score that is higher than the output-increasing efficiency score. The converse is true in the case of the decreasing returns to scale assumption as either AB or CD gets reduced and hence the output-increasing efficiency score is higher.

2.3.4. Peers of Firms and Slacks Associated with Inputs and Outputs

DEA is based on the assumption of convexity, which states that for any two points that are feasible, their convex combination is also feasible. This means that for two observed DMUs lying on the frontier one can prove that their convex combination is feasible and also lies on the frontier. Based on this assumption, DEA compares actual firms to virtual firms that are the weighted combinations of actual firms.

Peers or the so-called "composite set", are the firms that are on the frontier or on the best-performing practice frontier. These are used as the reference of comparison for inefficiently performing firms. In Figure 2.6, firms A, B and G lie completely to the northeast of the frontier and are inefficient. Firm A is projected onto the frontier and it falls on point D, which is an actual firm. Hence, firm D is the peer of firm A with respect to efficiency measurement. In this case, firm A is compared to the actual firm D. However, B' is the projection of firm B onto the frontier. Here, since it falls on the frontier between firms D and E, both D and E are the peers of firm B and firm B is compared to the weighted combination of D and E.

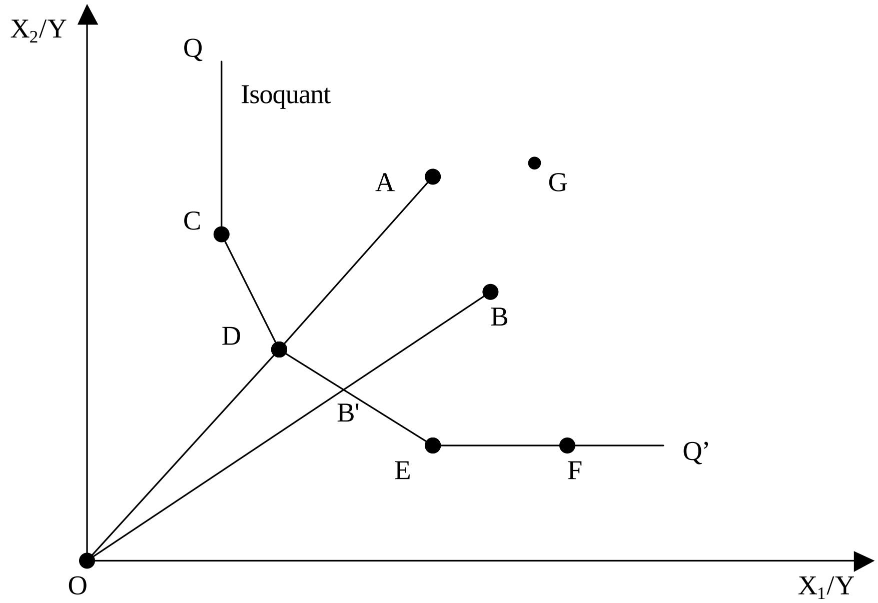


Figure 2. 6 Scatter Plot Representing Peers and Slack Inefficiencies

Efficiency measurement of operational units needs to consider the case when there are certain units or firms lying on that portion of the frontier that is parallel to one of the axes. As seen in Figure 2.6, the firm F lies on the frontier. So as per the discussions so far, firm F is rendered efficient as it lies on the frontier. However, it can be seen that firm F produces the same quantity of the output as E. It also uses the same quantity of input X_2 but more of X_1 for every unit of Y produced.

The amount of input X_1 given by the distance EF is the excess of input X_1 used by firm F. Hence the slack associated with input X_1 for firm F should also be taken into account while determining its efficiency score. A similar discussion pertains for the outputs, where the slacks with respect to the outputs would be termed as a shortfall in production.

2.3.5. Disposability of inputs and outputs

Disposability of inputs or outputs is the ability with which an input (output) can be disposed off holding the remaining inputs (outputs) constant while at the same time the resulting input (output) set still remains part of the production possibility set. The disposability concept of the inputs and the outputs is related to the definition of the production possibility set P. A production possibility set P, is the set of all points A (x, y) that satisfies the relationship: P = {(x, y): x can produce y}

Considering a single input and single output case as in Figure 2.7, the shaded portion is the production possibility set. Consider a production possibility set A(x, y) and its mapping into the production possibility set B (x, y_1) as shown.

$$A(x, y) \in P \Rightarrow B(x, y_1) \in P \; \forall \; B : y_1 \leq y \qquad \textbf{(2.3)}$$

The above condition defines a strong disposability of the outputs.

The weak disposability of the outputs is given by the condition in Equation 2.4. In a similar manner, it can be stated for the inputs.

$$A(x, y) \in P \Rightarrow C(x, y_1) \in P \; \forall \; C : y_1 \leq \mu y, 0 \leq \mu \leq 1 \qquad \textbf{(2.4)}$$

Fare et al. (1994) referred to disposability as the ability to stockpile or dispose of unwanted commodities. Thus the private disposal cost distinguishes the two different types of disposability. Strong disposability is the ability to dispose of an unwanted commodity with no private cost and weak disposability is the ability to dispose of an unwanted commodity at positive private cost. This concept will be explained in detail in the sections to come.

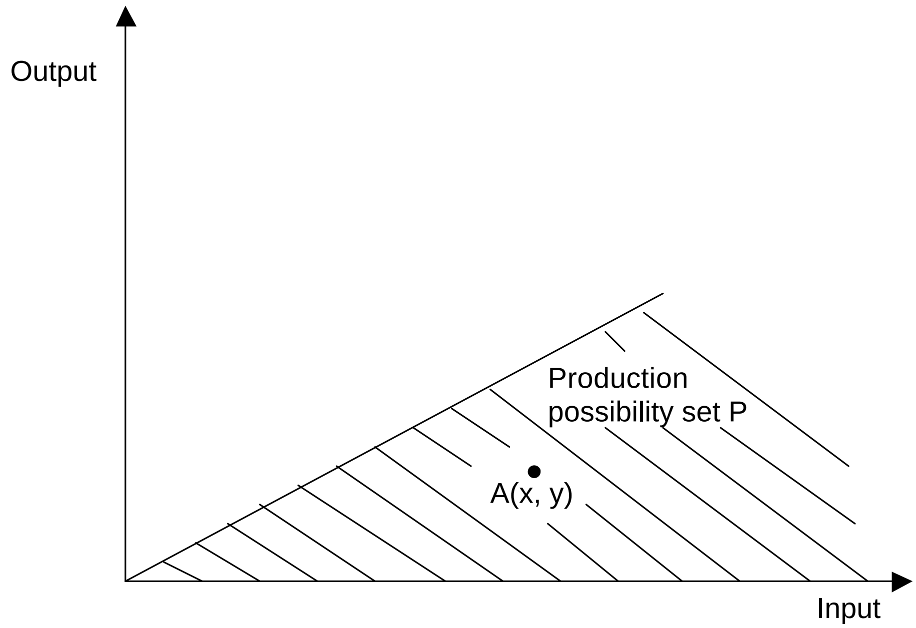


Figure 2. 7 Production Possibility Set

2.4. RELATIVE EFFICIENCY MEASUREMENT

The measurement of relative efficiency in the presence of multiple inputs and outputs was addressed by Farrell (1957) by assigning weights to the variables so that the overall relative efficiency score is actually a ratio of the weighted sum of the outputs to the weighted sum of the inputs.

$$\text{Efficiency} = \frac{\text{Weighted sum of outputs}}{\text{Weighted sum of inputs}} \qquad \textbf{(2.5)}$$

Thus the presence of multiple inputs and outputs are considered in the efficiency measurement process. Here, it is common to apply the same set of weights to the inputs or outputs of all firms. In this way, equal importance is given to a particular input or output for all the firms. Consider *n* firms with *m* inputs and *s* outputs. Let x_{ij} be the inputs and y_{rj} be the outputs of firm j. the mathematical representation of the above model would be as follows:

$$
\begin{aligned}
&\boldsymbol{h}_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \quad j = 1,2,3,\ldots n \\
&where, \\
&\boldsymbol{h}_j - efficiency\ of\ unit\ j \\
&u_r - weight\ on\ output\ r \\
&v_i - weight\ on\ input\ i \\
&y_{rj} - quantity\ of\ output\ r\ for\ unit\ j \\
&x_{ij} - quantity\ of\ input\ i\ for\ unit\ j
\end{aligned}
\qquad \textbf{(2.6)}
$$

The weights applied here have cost-price implications. The weights of the inputs correspond to their costs and the weights of the outputs to their prices. The model here results in efficiency scores in the range [0,1]. All the firms have a common set of weights assigned by the decision-maker. By assigning a common set of weights, the individual firms are not given the freedom to choose their own set of weights for their inputs and outputs. Thus the efficiencies of the firms are determined under this predefined set. Thus, in this case there is no possibility of increasing the efficiency score of a firm by way of assigning the weights that are most favorable for that firm. The next section explains the relaxation given to the DMU to choose its own weights, as well as the measurement process.

2.5. DATA ENVELOPMENT ANALYSIS

In the relative efficiency model, it is very difficult to find the cost of each input and the price of each output so that specific cost and price could be assigned to each input and output respectively. Further, the various firms organize their process differently hence value their inputs and outputs differently. This gives rise to differing weights. To overcome these drawbacks, Charnes, Cooper and Rhodes (1978) arrived at a mathematical programming approach that determines the weights and computes the efficiency score. The term Decision-Making Unit (DMU) will be used henceforth to represent a firm or an organization being evaluated.

2.5.1. The CCR Model

A fractional programming model known as the CCR model was developed by Charnes, Cooper and Rhodes (1978) to determine the efficiency score of each of the DMUs in a data set of comparable units. This model determines the best set of weights for each DMU when the problem is solved for each DMU under consideration.

The objective function maximizes the efficiency of the DMU using the weights u_r and v_i for the outputs and the inputs respectively. The weights are determined by the model such that the efficiency score of the DMU under consideration is maximum and when the same set of weights are applied to the other DMUs in the sample their efficiency score cannot exceed one. The mathematical formulation is provided below.

$$\begin{aligned}
&\max\ h_0 = \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} && \textit{for } DMU_0 \\
&\textit{subject to } \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1 && j = 1,2,3,\ldots n \\
&v_i \ge 0 && i = 1,2,3,\ldots m \\
&u_r \ge 0 && r = 1,2,3,\ldots s
\end{aligned} \tag{2.7}$$

The CCR model is a fractional program and it has to be converted into a linear program so that it can be solved easily. This is done by normalization i.e., the denominator of the objective function is equated to one and the first constraint corresponding to efficiency ratios of all DMUs in the sample is also modified as shown.

$$
\begin{aligned}
\max\ h_0 = \quad & \sum_{r=1}^{s} u_r y_{rj_0} && \qquad \textit{for } DMU_0 \\
\textit{subject to} \quad & \sum_{i=1}^{m} v_i x_{ij_0} \qquad\qquad = 1 && \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0 && \quad j = 1,2,3,\ldots n \\
& v_i \qquad\qquad\qquad\quad \ge e && \quad i = 1,2,3,\ldots m \\
& u_r \qquad\qquad\qquad\quad \ge e && \quad r = 1,2,3,\ldots s
\end{aligned}
\qquad \textbf{(2.8)}
$$

The above model is known as the multiplier model since it is developed from the fractional program and the input and output variables are multiplied with their respective weights. ε is a **weight restriction** that is introduced in the last two constraints. It is a lower bound or limit on the weights of the variables to make sure that none of the weights assigned to the input and output variables are zero. This assures that all the variables are taken into account while determining the efficiency score. In the absence of such a lower bound weight restriction, a DMU that is performing very badly with respect to one or more of the inputs or outputs could assign weights of zero to these variables and end up being efficient. By restricting these weights to be greater than ε, the decision-maker makes sure that none of the variables are neglected by any of the DMUs.

The linear program above is not unbounded because of the presence of the second constraint. In the absence of the second constraint, the linear program could typically assign increasing amount of weights to the outputs and continue to increase the value of the objective function. But this is constrained by the second constraint. Once the weights for the inputs are determined, the second constraint restricts the assigning of weights to the outputs, where the difference between the weighted sum of the outputs and the weighted sum of the inputs should be less than zero (0).

This model has *m+s* variables and *1+n+m+s* constraints. The dual would have *n+m+s+1* variables and *s+m* constraints. In general, *n* is quite large compared to *m+s* and hence the primal has a large number of constraints compared to the dual. Hence the primal is more difficult to solve. The dual of this linear program is obtained by assigning a variable to each constraint and transforming the constraints. The dual is known as the envelopment form, which is shown below.

$$
\begin{array}{llll}
\min & z_0 - \boldsymbol{e}\sum_{r=1}^{s} s_r^+ - \boldsymbol{e}\sum_{i=1}^{m} s_i^- & & \textit{for } DMU_0 \\
\textit{subject to} & x_{ij_0} z_0 & = \sum_{j=1}^{n} x_{ij}\,\boldsymbol{l}_j + s_i^- & i = 1,2,3,\ldots m \\
& \sum_{j=1}^{n} y_{rj}\,\boldsymbol{l}_j = y_{rj_0} + s_r^+ & & r = 1,2,3,\ldots s \\
& \boldsymbol{l}_j & \geq 0 & j = 1,2,3,\ldots n \\
& s_r^+ & \geq 0 & r = 1,2,3,\ldots s \\
& s_i^- & \geq 0 & i = 1,2,3,\ldots m \\
& z_0 & \textit{unconstrained} &
\end{array} \qquad \textbf{(2.9)}
$$

2.5.2. The BCC Model

The CCR model assumes constant returns to scale while determining the efficiency of the DMUs. Banker, Charnes and Cooper (1984) modified the CCR model by adding a constraint to account for the variable returns to scale. The difference between the two models is illustrated by Figure 2.8

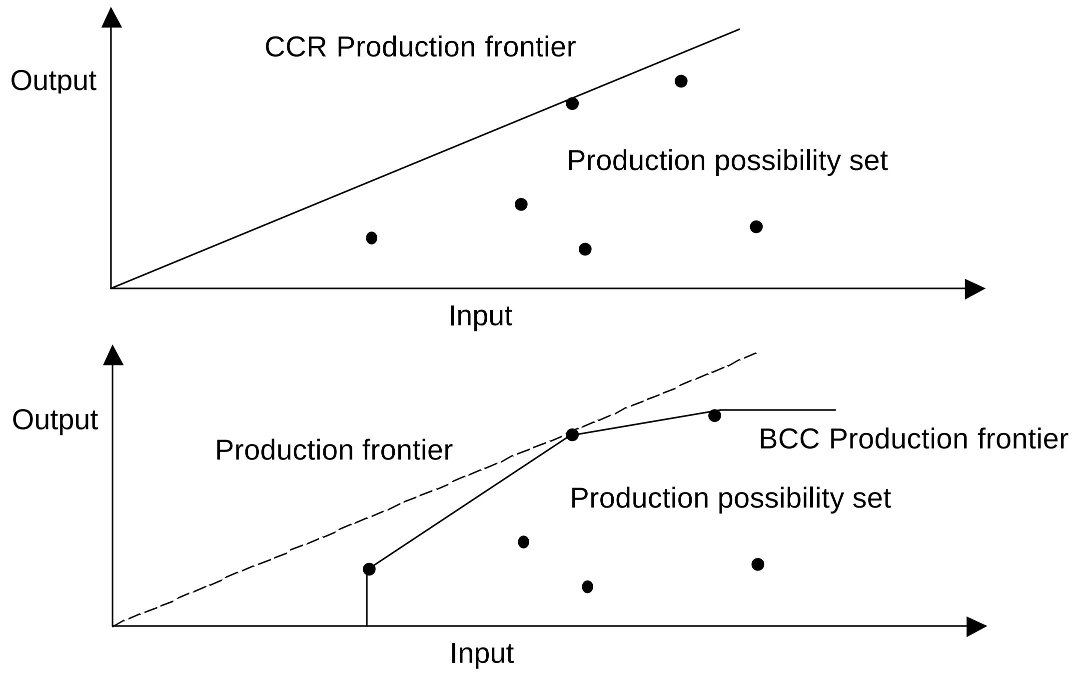


Figure 2. 8 The CCR and BCC Models

The envelopment form of the BCC model would be the same as the dual for the CCR model but with an additional constraint,

$$e\,l = \sum_{j=1}^{n} l_j = 1 \qquad \textbf{(2.10)}$$

2.6. NON-RADIAL EFFICIENCY MEASURES

So far, all the efficiency measures have determined the efficiency scores on a radial basis. In the case of the input-reducing models, the technical efficiency was determined by reducing all the inputs proportionally along a radial ray towards the origin from the point representing the firm or the DMU to a projected point on the frontier. In the case of the output-increasing models, a similar task is performed but by increasing the outputs along the ray away from the origin to a point on the frontier.

In the non-radial measure of technical efficiency, the comparison is done between the point representing the DMU and a point on the frontier that is not on the radial ray joining the origin and the point. Again in the case of the input-reducing models, the movement is in a direction so as to reduce one or more of the inputs and which is not radial. In the output-increasing models, the aim is to increase one or more of the outputs again in a non-radial direction. This can be explained with the help of the Figure 2.9.

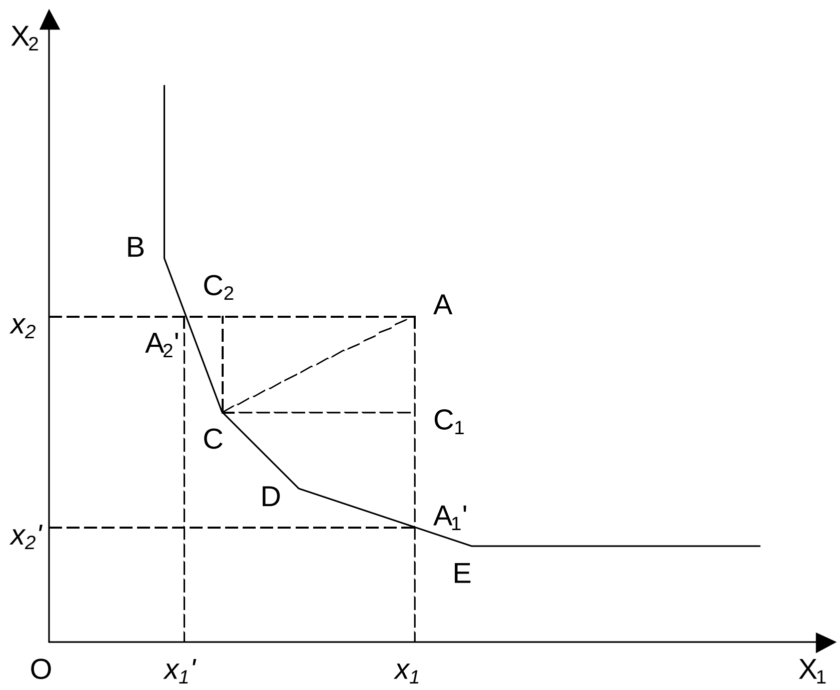


Figure 2. 9 Non-Radial Efficiency Measure

A is a DMU, which utilizes x_1 and x_2 units of the two inputs respectively. As seen from Figure 2.9, DMU A is inefficient, as it is not on the frontier. Unlike the earlier measure of technical efficiency, which would have considered a movement along the radial ray AO, the non-radial measure of technical efficiency, considers the movement along either AA_1' or AA_2'. In either of these cases, only one of the inputs gets reduced in the comparison. The other input is held constant. In a more than two input scenario (multidimensional) the equivalent would be to reduce one or more of the inputs in such a way that the resultant path is non-radial. This could be accomplished either by reducing a few of the inputs (but not all) proportionately or by reducing all of the inputs disproportionately.

In Figure 2.9, by holding X_1 constant and by reducing X_2, the point A_1' on the frontier is determined. The efficiency measure related to this observation on the frontier is x_2/x_2'. A similar argument can be made for holding input X_2 constant to arrive at the point A_2'.

Another non-radial measure of technical efficiency is possible where the movement is neither radial nor is parallel to one or more of the axes. The movement along AC illustrates this type of measure. The path is AC_1C or AC_2C. In either case, the input X_1 gets reduced by C_1C and input X_2 gets reduced by C_2C. It can be seen that there is disproportionate reduction of the two inputs along AC.

The **additive model** of Cooper et al. (1999) computes a different type of non-radial measure. This is illustrated in Figure 2.10, where the input-oriented and the output-oriented measures are combined to determine a score of technical efficiency. The following linear program illustrates the additive model. Here, the sum of all the slacks associated with each of the inputs and the outputs is maximized. Each DMU is projected to the farthest point on the frontier in a direction that tends to increase the output and reduce the input.

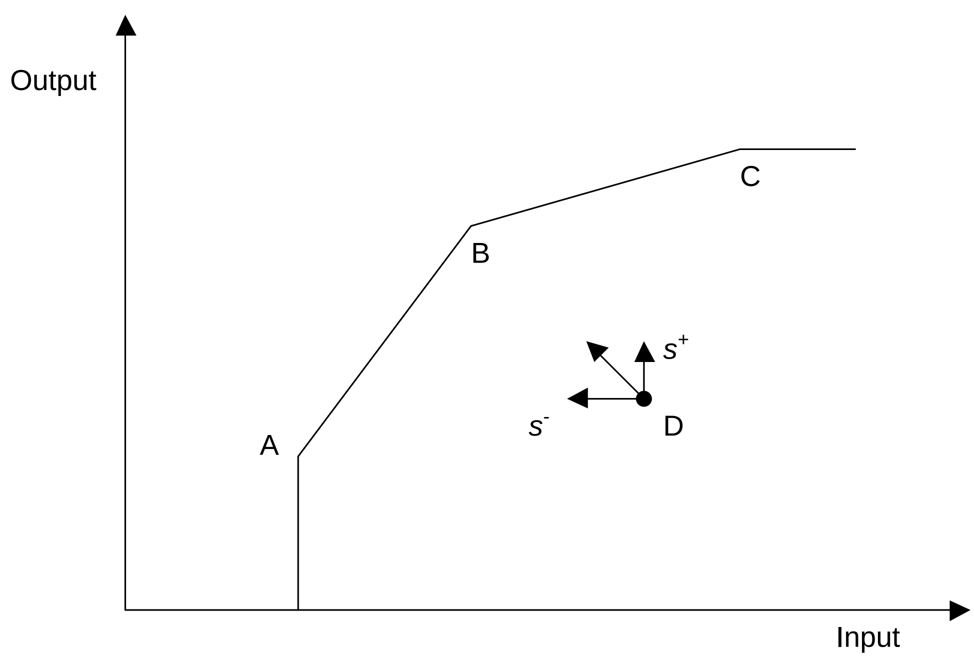


Figure 2. 10 The Additive Model

The mathematical formulation for the additive model is shown below.

$$
\begin{aligned}
&\max \quad z = es^{-} + es^{+} \\
&subjectto \quad X\lambda + s^{-} = x_0 \\
&Y\lambda - s^{+} = y_0 \\
&e\lambda = 1 \\
&\lambda \geq 0 \\
&s^{-} \geq 0 \\
&s^{+} \geq 0
\end{aligned}
\qquad \textbf{(2.11)}
$$

2.7. UNDESIRABLE OUTPUTS

So far, the discussion included the variables in the output set that were maximized. Nevertheless, this is not the case in the real world. There are outputs that are undesirable. Some examples of these outputs are pollution in industries, complications in operational procedures in hospitals and medical institutions, tax payments in financial firms, etc. These outputs are present in the output set along with the desirable outputs like units of production, number of operational procedures performed, number of customers, revenue generated, etc. Various techniques that consider the desirable and the undesirable outputs together in the modeling process are presented in this section. The premise is that while the desirable outputs are being maximized, the undesirable outputs should be minimized.

The main challenges that are considered in the modeling of the undesirable outputs are the following.

(i) One needs to consider the undesirable outputs in the modeling process along with the desirable outputs.

(ii) One requires that the undesirable outputs be reduced while the desirable outputs are increased.

An equivalent scenario of the undesirable outputs but on the input side is one where a subset of the inputs requires to be increased. For example, products that are recycled into the production process such as the plastic obtained in an injection molding firm, scrap, etc. These types of inputs need to be maximized mainly due to two reasons.

1. If output waste products are recycled, these can be fed back into the transformation process. By using these products as inputs as much as possible, problems of disposal and environmental concerns that are associated with these types of outputs can be avoided. This helps the firms to reduce the amount of waste products (undesirable outputs) that are disposed of into the environment and helps firms to meet environmental regulatory conditions.

2. By maximizing the usage of these inputs, the usage of the other inputs could be reduced through substitution. In this way, production costs could be reduced.

Scheel (2001) introduced various techniques to address the challenges of incorporating the undesirable outputs into the DEA model, minimizing the undesirable while maximizing the desirable outputs. He introduced new radial measures, which assume that any change of the output level will involve both the desirable and the undesirable outputs.

Scheel defines the technology set as

$$T = \{(x, y) : \mathbf{l}^T X \le x,\ \mathbf{l}^T Y \ge y,\ \mathbf{l} \ge 0,\ \mathbf{l}^T e = 1\} \qquad \textbf{(2.12)}$$
$$\textit{where} \quad e = (1, \ldots 1)^T$$

The next section addresses two known approaches to include undesirable outputs in the modeling process.

2.7.1. Efficiency Classifications

There are various approaches to incorporate undesirable outputs into the DEA model. They are classified as indirect and direct approaches.

1. **Indirect approaches** transform the values of the undesirable output variables by a monotone decreasing function so that they can be included in the model along with the desirable outputs in the technology set T and are maximized. In this way, by maximizing the transformed values, the original undesirable output values are minimized.

2. **Direct approaches** on the other hand, include the undesirable output data directly into the DEA model but instead modify the assumptions of the model in order to consider the undesirable outputs appropriately.

Scheel considers *X* to be the matrix of the inputs, *Q* the matrix of undesirable outputs and *P* the matrix of desirable outputs. He assumes that the matrices *X, Q* and *P* are non-singular matrices with no vanishing rows or columns (determinant is not zero). By this assumption, we mean that the set of input variables are linearly independent and so are the set of desirable and undesirable outputs.

The first indirect approach of incorporating the undesirable outputs into the model is by transforming them by the additive inverse method. This involves including the undesirable outputs using values *f (Q)=-Q.* The author refers to this approach as [ADD] and it was suggested by Koopmans (1951). The technology set defined by this approach is the same as the one defined by the approach [INP]. [INP] is an approach where the undesirable outputs *Q* are considered as inputs and are considered together with the other inputs. Thus in the process of reducing the inputs, the undesirable outputs are also reduced.

Another approach is the [TRβ] in the sense of Ali and Seiford (1990) where a large scalar β is added to each of the undesirable output values such that the transformed values are positive. The transformation is done using

$$f_r^{\,j}(Q) = -q_r^j + \boldsymbol{b}_r \qquad \textbf{(2.13)}$$

Here the transformation is done for the r^{th} undesirable output of the j^{th} firm. The last of the indirect approaches is the multiplicative inverse [MLT] suggested by Golany and Roll (1989) where the transformation is done by using the function

$$f_r^{\,j}(Q) = 1/u_r^j \qquad \textbf{(2.14)}$$

Then the data of the undesirable outputs is included with the data of the desirable outputs.

There is also a relation between [MLT] and [ADD]. If a DMU is efficient using [MLT], then it is also efficient when using [ADD] (Scheel (2001)). Hence, as shown by the author, it is clear that [MLT] is a more “restrictive” approach than [ADD] and thus it is more difficult to become efficient under [MLT] than under [ADD].

Some of the problems and issues associated with the above approaches are summarized here. By using the [INP] approach and by considering a part of the output set variables as inputs, the basic input-output structure of the transformation process is modified. This approach is slightly different from the other approaches in the sense that, in the [ADD], [MLT] and [TRβ], the transformed data remains as a part of the output set. But in the case of [INP] the transformation as such takes the data and considers them as inputs. Also, in the case of the [MLT] formulation, because of the transformation done to the data set, we run into problems associated with convexity and non-linearity. In the case of the indirect approaches, the framework assumes that the transformed data has its own meaning, like the undesirable output “mortality rate” and its additive inverse “survival rate”. This may not be the case with all processes. In many real life applications, the transformation of the data may not make sense. This motivates the researcher to explore other direct approaches where no transformation is done to the data set, but they are used with necessary modifications in the modeling assumptions.

Fare et al. (1986) was the first to apply output oriented DEA analysis to steam electric plants. In this paper the authors defined radial efficiency measures for equi-proportional increase of all the outputs, both desirable and undesirable. The technology set was characterized by strong and weak disposability of the undesirable outputs in order to check for production congestion that is explained later in this chapter under the section Congestion.

Fare et al. (2000) provide a formal index number of environmental performance that can be computed using DEA techniques. The procedure is to determine an output distance function based on the degree to which the desirable outputs could be expanded for each DMU. This is represented by the quantity index of desirable outputs. . For the undesirable outputs, the degree of reduction possible is determined for each DMU that is represented by the quantity index of undesirable outputs. The overall environmental performance index is given by the ratio of the quantity index of the desirable outputs to the quantity index of the undesirable outputs.

Let us define the following notation for the vector of inputs, the desirable outputs and the undesirable outputs respectively.

$$\begin{aligned} x &= (x_1, x_2, \dots x_M) \in \Re_+^M \\ p &= (p_1, p_2, \dots p_N) \in \Re_+^N \\ q &= (q_1, q_2, \dots q_J) \in \Re R_+^R \\ y &= (p, q) \in \Re_+^{N+R} \end{aligned} \qquad \textbf{(2.15)}$$

Hence, the technology has all feasible vectors *(x,y)*

$$T = \{(x, y) : x \text{ can produce } y \text{ where } y = (p, q)\} \qquad \textbf{(2.16)}$$

The various assumptions introduced are as follows.

1. Weak disposability of outputs:

$$If\ (x, y) \in T \text{ and } 0 \le \boldsymbol{q} \le 1 \text{ then } (x, \boldsymbol{q}y) \in T \text{ where } y = (p, q) \qquad \textbf{(2.17)}$$

The weak disposability assumption makes sure that both the desirable and the undesirable outputs can be disposed proportionally. It also implies that it is not possible to reduce only the undesirable outputs holding the inputs and the desirable outputs constant.

According to the null joint assumption discussed below, it is technically impossible to produce only desirable outputs without producing any undesirable outputs. Moreover, to totally eliminate the undesirable outputs, the desirable outputs cannot be produced.

2. Null jointness

This assumption means that it is technically impossible to produce only desirable outputs without producing any of undesirable outputs. It also means that the only way to totally eliminate the production of the undesirable outputs is to stop the production of the desirable outputs.

$$If\ (x, y) \in T\ where\ y = (p, q)\ and\ q = 0\ then\ p = 0 \qquad \textbf{(2.18)}$$

3. *T* is a closed set. This assumption implies that the production possibility set also includes all the points on the frontier. Any point on the frontier that is a vertex represents a real DMU. Other points on the frontier are virtual composite DMUs used for comparison after the projection onto the frontier.
4. Inputs are freely or strongly disposable. This means that, if an amount *y* can be produced from *x*, then *y* can be produced from any *x'* ≥ *x*. In words, it means that the amount of input can be increased without an increase in the amount of outputs.
5. The desirable outputs are freely or strongly disposable. This means that if a given quantity of desirable output *p* can be produced from *x*, and for any amount *p'* ≤ *p*, *p'* can also be produced from *x*.

Then the authors define an output distance function for the desirable outputs.

$$D_p(x, p, q) = \inf\{\boldsymbol{q} : (x, p/\boldsymbol{q}, q) \in T\} \qquad \textbf{(2.19)}$$

Let x^0 and q^0 be the input and the undesirable output vectors and let p^k and p^l be the desirable output vectors that are being compared. Then the quantity index of good outputs is given by

$$Q_p(x^0, q^0, p^k, p^l) = D_p(x^0, p^k, q^0) / D_p(x^0, p^l, q^0) \qquad \textbf{(2.20)}$$

This index satisfies the following properties.

1. Homogeneity:

$$Q_p(x^0, q^0, \boldsymbol{l}p^k, p^l) = \boldsymbol{l}Q_p(x^0, q^0, p^k, p^l) \qquad \textbf{(2.21)}$$

2. Time reversal:

$$Q_p(x^0, q^0, p^k, p^l).Q_p(x^0, q^0, p^l, p^k) = 1 \qquad \textbf{(2.22)}$$

3. Transitivity:

$$Q_p(x^0, q^0, p^k, p^l).Q_p(x^0, q^0, p^l, p^s) = Q_p(x^0, q^0, p^l, p^s) \qquad \textbf{(2.23)}$$

4. Dimensionality:

$$Q_p(x^0, q^0, \lambda p^k, \lambda p^l) = Q_p(x^0, q^0, p^k, p^l) \qquad \textbf{(2.24)}$$

A distance function is also determined for the undesirable outputs similar to the desirable outputs. It is given by:

$$D_q(x, p, q) = \sup\{\lambda : (x, p, q/\lambda) \in T \qquad \textbf{(2.25)}$$

Then the undesirable output quantity index is given by:

$$Q_b(x^0, y^0, b^k, b^l) = D_b(x^0, y^0, b^k) / D_b(x^0, y^0, b^l) \qquad \textbf{(2.26)}$$

The environmental performance index is defined as:

$$E^{k,l}(x^0, p^0, q^0, p^k, p^l, q^k, q^l) = \frac{Q_p(x^0, q^0, p^k, p^l)}{Q_q(x^0, p^0, q^k, q^l)} \qquad \textbf{(2.27)}$$

The linear programs used to compute the values of the distance functions are as follows.
<u>Model L2</u>: Desirable outputs:

$$\begin{aligned}
&(D_p(x^0, p^{k'}, q^0))^{-1} = \max \theta \\
&s.t. \quad \sum_{k=1}^{K} z_k p_n^k \geq \theta p_n^{k'}, n = 1,...N \\
&\sum_{k=1}^{K} z_k q_r^k = q_r^0, r = 1,...R \\
&\sum_{k=1}^{K} z_k x_m^k \leq x_m^0, m = 1,...M \\
&z_k \geq 0, k = 1,...K
\end{aligned} \qquad \textbf{(2.28)}$$

$Q_p(x^0,q^0,p^k,p^l)$ is the ratio of the distance functions for desirable outputs of the DMU under consideration and the reference DMU. Each DMU is taken into consideration and the distance function for the desirable outputs is calculated using the above linear program which is the numerator of the ratio $Q_p(x^0,q^0,p^k,p^l)$. The denominator, which is the distance function for the reference DMU is computed.

Similarly, the following linear program calculates the distance function for undesirable outputs:

$$\begin{aligned}
&(D_q(x^0, p^0, q^{k'}))^{-1} = \max\ l \\
&s.t.\quad \sum_{k=1}^{K} z_k p_n^k \geq p_n^0, n = 1,\ldots N \\
&\sum_{k=1}^{K} z_k q_r^k = l q_r^{k'}, r = 1,\ldots R \\
&\sum_{k=1}^{K} z_k x_m^k \leq x_m^0, m = 1,\ldots M \\
&z_k \geq 0, k = 1,\ldots K
\end{aligned} \tag{2.29}$$

Fare et al. (2000) apply this technique to measure the environmental performance of 19 countries with the following variables.

S. No.	Name of variable	Type of variable	Unit
1	Oil consumption	Input	Million tons of equivalent
2	Labor	Input	Number of workers in thousands
3	Capital stock	Input	Trillions of US $
4	Gross Domestic Production	Desirable output	Billions of US $
5	Carbon dioxide emissions	Undesirable output	Millions of tons
6	Nitrogen oxide emissions	Undesirable output	Thousands of tons
7	Sulfur dioxide emissions	Undesirable output	Thousands of tons

Table 2.1 - Variables used by Fare, et al. (2000)

The authors then conduct a set of non-parametric tests to determine if there is a relationship between the overall environmental indicator and the country characteristics. For this purpose, the countries are split into two groups. Group A comprises of those countries with an environmental performance index below the median indicator and group B consists of those with an index equal to or better than the median. Then they determine if there is a statistically significant difference between the two groups with respect to GDP/Population and Oil consumption/GDP.

The authors found that there is no significant statistical difference between the two groups of countries with respect to the per capita GDP. In previous studies, the pollutants have been considered separately in the analysis and it has been found that, CO_2 has a positive relationship with GDP per capita and the other pollutants have a negative relationship. Based on the work of Fare et al., there seems to be no clear-cut relationship between the emissions CO_2, SO_x, NO_x and the per capita GDP. In this case, all three pollutants are considered together in computing the performance index. Also, in case where only CO_2 and NO_x were considered and in case where only CO_2 and SO_x were considered, the authors did not find any significant correlation with GDP. Their indices account for the production of goods and the use of inputs simultaneously. This is achieved by the use of distance functions. In addition, since the data used is just for a single year, it fails to reveal the time series relationships.

If data over time is available, the panel nature could be exploited to compute environmental performance indices over time to determine the performance of countries over a period of time. This would also mean that no single country need to be considered as the reference to calculate the desirable and undesirable output quantity indices. Each country can be evaluated based on its performance over a period of time.

As it can be seen, in the case of the index number approach, the performance of the DMUs are evaluated separately based on the desirable outputs and the undesirable outputs and then an index is computed. This model cannot handle a combined structure for improvements in the desirable outputs and the undesirable outputs in the same model. The improvements in the desirable and the undesirable outputs are linear and independent of one another.

2.7.2. Hyperbolic Efficiency Measure

Fare et al. (1989) introduced a different approach to incorporate both the desirable outputs and the undesirable outputs in the model. They allowed the desirable outputs to increase by some proportion and at the same time allowed the undesirable outputs to decrease by the same proportion. The result was a northwesterly hyperbolic path as shown in the Figure 2.11 below.

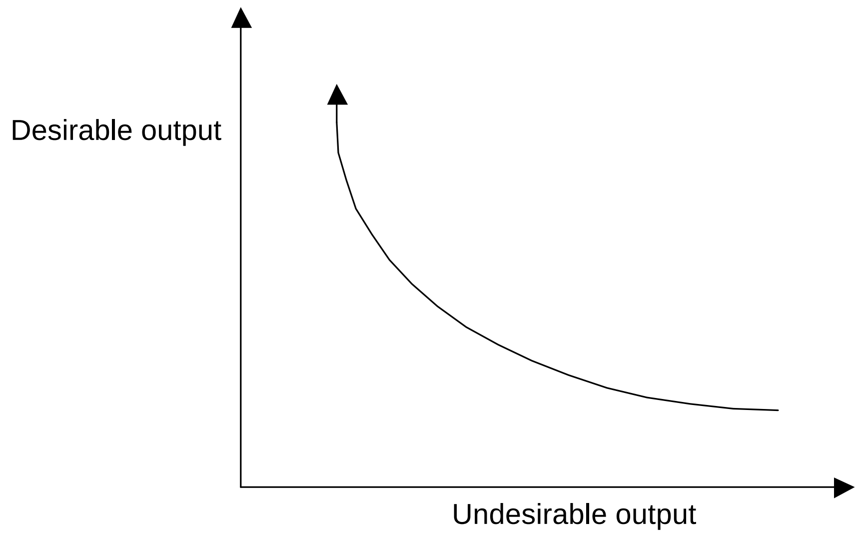


Figure 2. 11 Hyperbolic Path

The main difference in the measurement between the radial and the hyperbolic efficiency measures is that in the case of the radial efficiency measure, the target of the performance score is to determine the extent to which a firm is technically efficient. Increasing both the desirable output and the undesirable output renders a greater technical efficiency but, due to the increase in the undesirable output, it is not environment-friendly. However, in the case of the hyperbolic path, the efficiency measure makes sure that as the desirable output increases, the undesirable output decreases. This is incorporated into the model in such a way that both the outputs change by the same proportion but in different directions.

Another interesting fact to be noted from the hyperbolic approach in Fare et al. (1989) is that, for greater quantities of undesirable outputs, small displacements to the frontier lead to a greater unit decrease in quantity of the undesirable output than the unit increase in the desirable output. This can clearly be seen from Figure 2.11 as the initial part of the hyperbolic path has a slope tending to 0° and it gradually increases until it reaches 90°. As the quantity of the undesirable output decreases, the initial part of the hyperbolic path becomes more and more steep. This results in a tremendous increase in the desirable output right from the start of the hyperbolic path as depicted by Figure 2.12 below.

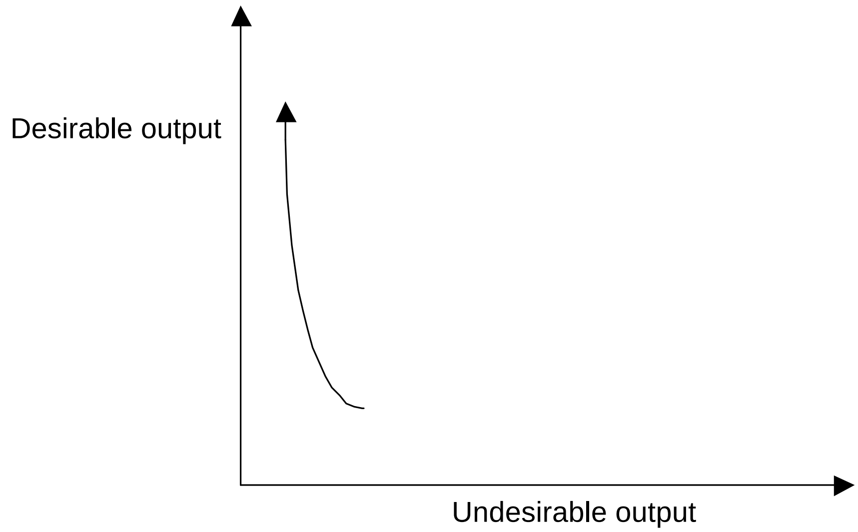


Figure 2. 12 Hyperbolic Path for a Firm with a Low Quantity of Undesirable Output

Implementing the hyperbolic approach, Zofio and Prieto (2001) assess the environmental performance of a set of producers by grading their ability to produce "the largest equi-proportional increase in the desirable output and decrease in the undesirable output."

The authors assume that the firms *k={1,…K}* use a set of *x* inputs to produce *y* outputs out of which *p* are desirable and *q* are undesirable.

$$\begin{aligned} x &= (x_1, x_2, \dots x_M) \in \Re_+^M \\ p &= (p_1, p_2, \dots p_N) \in \Re_+^N \\ q &= (q_1, q_2, \dots q_R) \in \Re_+^R \\ y &= (p, q) \in \Re_+^{N+R} \end{aligned} \qquad \textbf{(2.30)}$$

The reference technology is modeled in the following manner.

$$R : \Re_+^M \to R(x) \subseteq \Re_+^{N+R} \qquad \textbf{(2.31)}$$

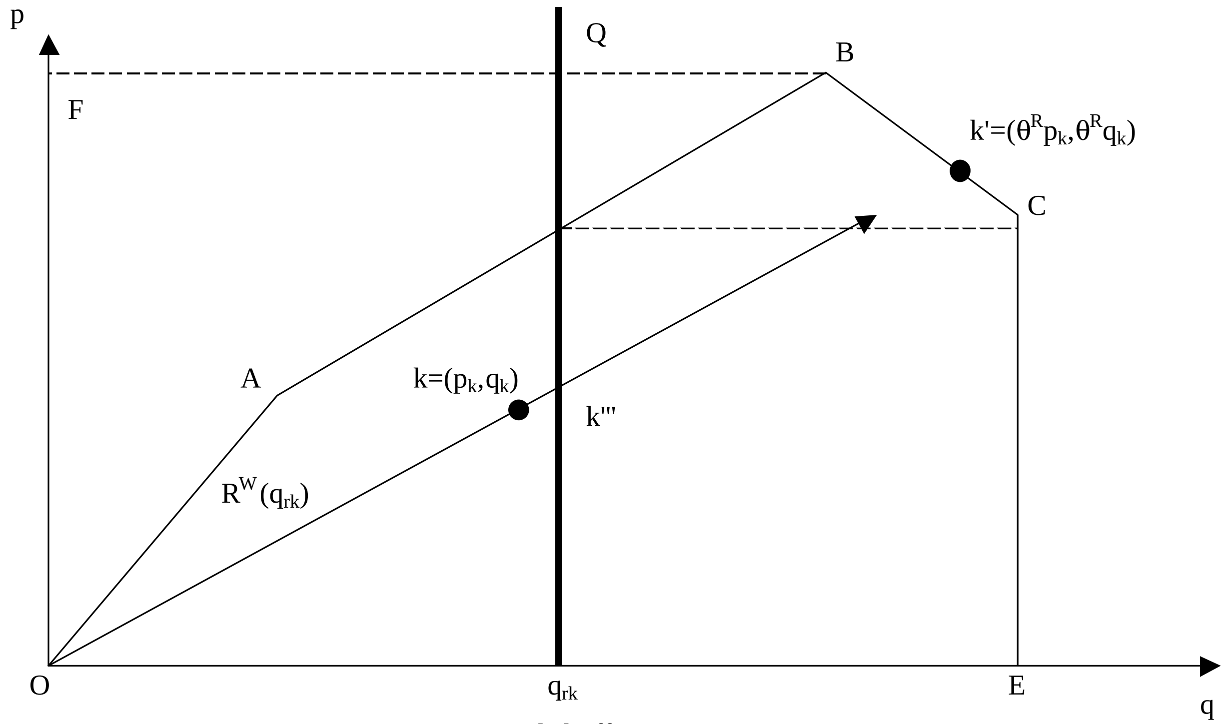


Figure 2. 13 Radial Efficiency Measure

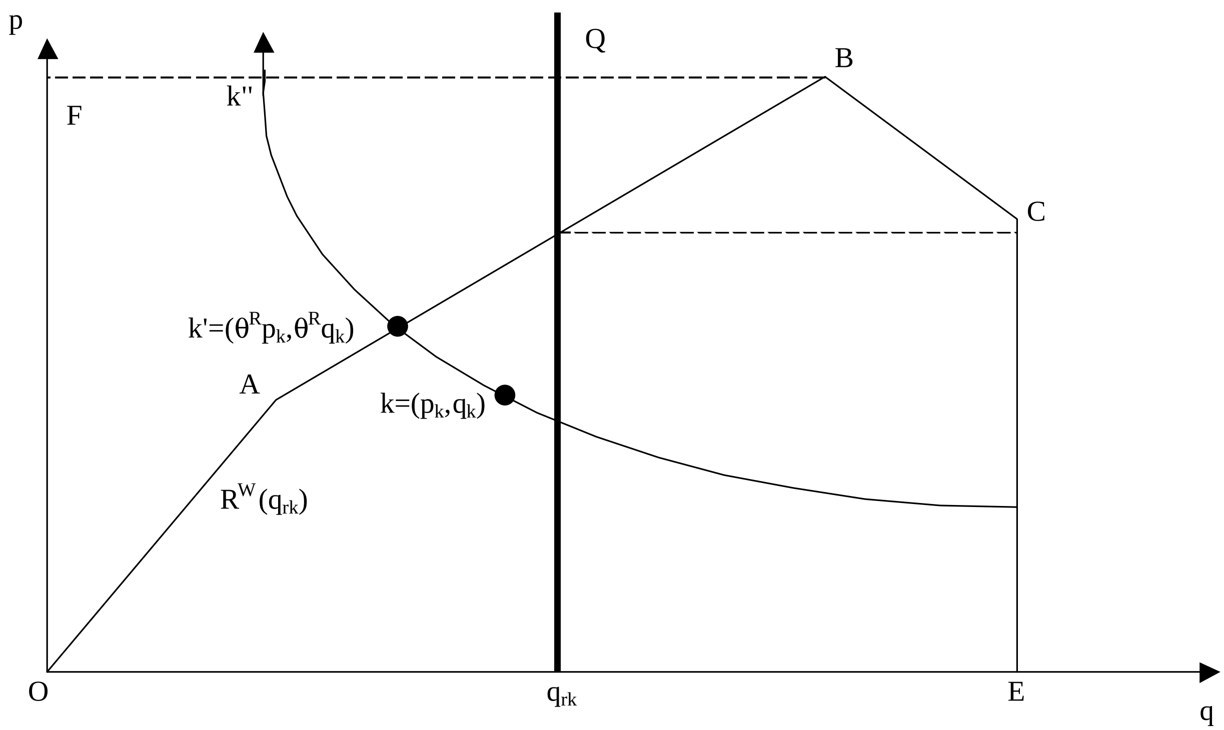


Figure 2. 14 Hyperbolic Efficiency Measure

Here the strong disposal reference technology for the inputs and the outputs is represented by the output set: which is represented by OFBCE in Figures 2.14 and 2.15.

$$R^S(x):[(p,q): p \leq Pz, q \leq Qz, x \geq Xz; z \in \Re_+^K] \qquad \textbf{(2.32)}$$

In the above expression, the superscript in $R^S(x)$, represents strong disposability. $\Re^K_+$ is the standard notation to represent the positive side of a K-dimensional space.

The figures 2.14 and 2.15 graphically illustrate the existence of weak disposability of the undesirable outputs. Considering the section OAB of the frontier in either of the figures, it can be seen that any reduction in the undesirable outputs results in a reduction of the desirable outputs. If the reduction in the desirable outputs has to be avoided, then the desirable outputs have to be held constant. This results in an increase in the usage of the inputs. In either case there is a cost associated with the disposal of the undesirable outputs and it is not free, hence the weak disposability.

The weak disposal reference technology for the undesirable outputs is the output set, which is represented by OABCE in Figures 2.13 and 2.14.

$$R^W(x):[(p,q): p \leq Pz, q = Qz, x \geq Xz; z \in \Re_+^K] \qquad \textbf{(2.33)}$$

Considering the same production possibility set, the functional representations of the reference technologies are given by an output distance function provided by Shephard (1953).

$$D^R(p,q) = \{\boldsymbol{q}^R : y\boldsymbol{q}^R \in R(x)\} \qquad \textbf{(2.34)}$$

A similar hyperbolic distance function is given as follows.

$$D^H(p,q) = \{\boldsymbol{q}^H : (p\boldsymbol{q}^H, q/\boldsymbol{q}^H) \in R(x)\} \qquad \textbf{(2.35)}$$

These distance functions are regarded as performance measures and are evaluated to determine the scores of the firms. The final efficiency score is given as the reciprocal of the distance function. If the expansion of either the radial or the hyperbolic distance functions is infeasible, then the respective distance function has a value of one, i.e. $q^R = q^H = 1$.

Such firms are characterized as efficient. This happens for the DMUs that are on the frontier. On the other hand, if the distance functions are expandable, then $q > 1$ and hence the production process is considered inefficient.

The radial and the hyperbolic measures of the above distance functions can now be defined as below.

$$\begin{aligned} E_{1k}^{R}(p_k, q_k, x_k) &= \max[\boldsymbol{q}^R : (\boldsymbol{q}^R p_k, \boldsymbol{q}^R q_k) \in R^S(x_k)] \\ E_{1k}^{H}(p_k, q_k, x_k) &= \max[\boldsymbol{q}^H : (\boldsymbol{q}^H p_k, q_k / \boldsymbol{q}^H) \in R^S(x_k)] \end{aligned} \quad \textbf{(2.36)}$$

These radial measure is computed using the formulation below in 2.37. The hyperbolic measure is computed using 2.38, which is explained later in the next section on the next page.

$$\begin{aligned} & E_1^S(p_k, q_k, x_k) = \max[\boldsymbol{q} : (\boldsymbol{q}p_k, \boldsymbol{q}q_k) \in R^S(x_k)] \\ & s.t. \\ & \boldsymbol{q}p_k \le Pz \\ & \boldsymbol{q}q_k \le Qz \\ & Xz \le x_k \\ & z \in \Re_+^K \end{aligned} \quad \textbf{(2.37)}$$

As it can be seen, the radial measure tries to increase both the desirable and the undesirable outputs along the radial direction away from the origin. This assumes that the undesirable output cannot be decreased simultaneously while the desirable outputs are being increased.

Brannlund et al. (1995) studied radial efficiency levels in the Swedish pulp and paper industry with environmental standards. The authors used a ratio of the restricted profit function and the unrestricted one to determine the effect of the regulation. But the efficiency and the productivity gains are determined with the help of ray vectors that increase both the desirable as well as the undesirable outputs. Obviously, this approach does not seem reasonable in cases where environmental concerns are central to the analysis. The hyperbolic approach takes care of this problem, since it considers the firms in the reference sets that not only have increased amount of desirable outputs but also have a reduced amount of undesirable outputs. The hyperbolic approach thus can be applied to processes where the environmental concerns are central to the problem. In the next section, the concept of congestion is introduced and the method to determine the presence of any congestion in a production process is discussed.

2.7.3. Congestion

A production process is said to exhibit strong disposability of undesirable outputs if the undesirable output can be freely disposed without any change in the desirable outputs or in the inputs. The weak disposability of undesirable outputs implies that a reduction in the undesirable output forces a proportional reduction in the desirable output as well. This leads to congestion in production.

The existence of congestion can be determined by evaluating the firm successively under strong and weak disposability assumptions. If the efficiency scores obtained from both the assumptions are the same, then the firm's production process is not affected by congestion. Hence the undesirable outputs can be reduced without any reduction in the desirable outputs. On the other hand, if the scores are different, then congestion is said to exist and weak disposability is binding. Any reduction in the undesirable output is associated with a loss in the desirable output. This loss can be determined by computing both the scores and comparing them.

We assume the strong disposal reference technology $R^S(x)$ and solve the following linear program to determine the efficiency score.

Model L1:

$$
\begin{aligned}
&E_1^S(p_k, q_k, x_k) = \max[\theta : (\theta p_k, \theta^{-1} q_k) \in R^S(x_k)] \\
&s.t. \\
&\theta p_k \le Pz \\
&q_k / \theta \le Qz \\
&Xz \le x_k \\
&z \in \Re_+^K
\end{aligned}
\qquad \textbf{(2.38)}
$$

In the above linear program, the constraints for the desirable and the undesirable outputs have a '≤' which accounts for the strong disposability of both the outputs. Since the above program is nonlinear in *q*, the problem cannot be solved by linear programming techniques. Fare et al. (1989) proposed the following expression, which is linear in *q* as a linear approximation to the nonlinear constraint. The assumption behind this is that *q* can take a value greater than 1 (one) and it converges at *q = 1*.

$$2q_k - \boldsymbol{q} q_k \leq Qz \tag{2.39}$$

This approach has some drawbacks. The linear approximation and the nonlinear constraint are equal only for *q = 1*. Also as the value of *q* diverges from one, the linear approximation diverges from the actual nonlinear constraint and hence the approximation error increases.

An alternate transformation that works around this problem is given by

$$(q_{tk} / \boldsymbol{q} \leq \sum_{k=1}^{K} q_{tk} z_k)^{-1} \tag{2.40}$$

This can be equivalently expressed as the inequality below, which is linear in θ

$$\boldsymbol{q}(q_{tk})^{-1} \geq (\sum_{k=1}^{K} q_{tk} z_k)^{-1} \tag{2.41}$$

The undesirable output weak disposability reference technology is given by $R^W(x)$. The corresponding efficiency score is determined by the following linear program.

$$\begin{aligned}
&E_1^W(p_k, q_k, x_k) = \max[\boldsymbol{q} : (\boldsymbol{q} p_k, \boldsymbol{q}^{-1} q_k) \in R^W(x_k)] \\
&s.t. \\
&\boldsymbol{q} p_k \leq Pz \\
&q_k / \boldsymbol{q} = Qz \\
&Xz \leq x_k \\
&z \in \Re_+^K
\end{aligned} \tag{2.42}$$

As it can be seen, ‘=’ is used in the undesirable output constraint to account for the weak disposability of these undesirable outputs while still holding the desirable outputs as strongly disposable.

Then the ratio E_{1k}^S/E_{1k}^W is defined to check for production congestion. If the ratio equals one, then the production does not present weak disposability of undesirable outputs. If the ratio is greater than one, then weak disposability congests the production process.

Exhibiting strong or weak disposability is based on the technology of the firm and it is internal to each firm. However, the regulatory standard that is imposed on the firm is external to the firm. It is to be noted that, the extent and the proportion of strong or weak disposability is different for each firm but the regulatory standard is the same for all the firms under comparison.

2.7.4. Production possibilities and environmental standards

The following model introduces quantitative constraints to incorporate the environmental standard.

$$
\begin{aligned}
&E_r(p_k, q_k, q_{rk}, x_k) = \max[\theta : (\theta p_k, \theta^{-1} q_k) \in R^S(q_{rk})] \\
&s.t. \\
&\theta p_k \le Pz \\
&q_k / \theta \le Qz \\
&Xz \le x_k \\
&Qz \le q_{rk} \\
&z \in \Re_+^K
\end{aligned}
\qquad \textbf{(2.43)}
$$

The linear program 2.43 has the same technology set as 2.38 except for the introduction of the new constraint to account for the environmental standard. Q_Z is the undesirable output of the composite unit that is used for the comparison. The new constraint implies that the undesirable output of the composite unit has to satisfy the environmental standard set, q_{rk}

The following ratio is used to check if the environmental standard affects the production process *k*.

$$E_{1k}^{S} / E_{rk} \qquad \textbf{(2.44)}$$

The numerator is obtained from 2.38 while the denominator is obtained from 2.43.

If the ratio is greater than one, then the standard binds the production while if its less than one, no costs are associated with the regulation.

Thereupon the authors Zofio and Prieto (2001) determine the limits relating to the undesirable outputs and the production. For a given piecewise reference technology, they determine the minimum amount of undesirable output that is required to carry out the production. Also, they compute the minimum amount of production of the undesirable output that is possible without sustaining output congestion.

This methodology is applied to manufacturing industries in 14 OECD countries where the undesirable output is CO_2. In the formulation of the model, it is assumed that all the countries have access to the efficient production frontier. In their case, this seems as a reasonable assumption since the data consists of only the most advanced OECD countries.

The index approach Fare et al. (2000) that has been discussed previously does not discuss the congestion involved with the technology by making the weak disposability assumption of the undesirable outputs. Further, the model does not consider any environmental policies. The model by Zofio and Prieto (2001) however addresses these issues.

The models by Fare et al. (2000) and Zofio and Prieto (2001) have to be modified to facilitate the incorporation of the strategic policies or measures that originate from the top-level management or from the individual DMUs. In such cases, when one wants to look at the performance of the DMUs with priorities being assigned, the model formulation needs to have a preferential structure built into it. Then the weights included in the model determine the relative importance given to the increase of the desirable outputs and the decrease of the undesirable outputs.

In the case of the Zofio and Prieto (2001) approach, it's a hyperbolic measure where the same factor is applied for both the increase of the desirable outputs and the decrease of the undesirable outputs. Still, the model fails to handle improvements of the two different types of outputs and the inputs with a preferential structure.

2.7.5. Non – DEA Approach – Pair wise Dominance

According to Otis (1999) a pair wise dominance approach is important for frontier models. This dominance production plan according to Koopmans (1951) is a production plan that uses less of inputs for the same amount of output or produces more of outputs for the same level of inputs. Figure 2.12 illustrates the two input production system with the outputs being fixed. The four sets are shown in the figure. Strong dominance refers to a production plan that has all inputs less than the reference while weak dominance means that at least one input is less than the reference and this is represented by the dotted lines.

In the figure below, the Dominated set would consist of all DMUs that consume more of the two inputs than the reference DMU and produce the same amount of outputs. On the other hand, the Dominating set would comprise of all DMUs that consume lesser of the two inputs and produce the same amount of the outputs. The Dominance Indifferent set will have DMUs that consume more of one of the two inputs and lesser of the other input.

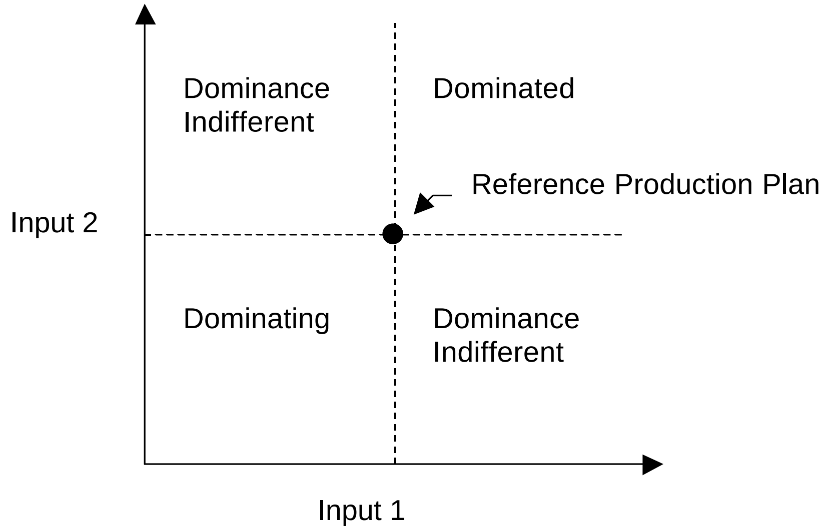


Figure 2. 15 Classification based on Dominance

By the inclusion of undesirable outputs in the pair wise dominance approach, the following news sets are defined.

1. Technically Dominating-Preferred Environmentally
2. Technically Dominated Non-Preferred Environmentally
3. Technically and Environmentally Dominance Indifferent

These sets can be explained in a similar manner with respect to the desirable and the undesirable outputs.

It can be seen that the above requires that a production plan not only represents an increase in the production efficiency but at the same time must also not represent an increase in the undesirable outputs to be considered as a part of the dominating set.

There are instances in the process of efficiency measurement where the decision-maker would like to give preference for the improvement of certain inputs and/or outputs and/or to hold certain others constant. These requirements tend to give rise to multiple objectives and a modified framework is necessary to handle these kinds of requirements. In the next section, the concept of goal programming is discussed and the way in which it is used in DEA to handle multiple objectives. This concept is termed as GoDEA and was first used by Charnes et al. (1988) to determine the performance of individual DMUs using some ‘ideal’ input/outputs as targets.

2.8. GOAL PROGRAMMING AND DATA ENVELOPMENT ANALYSIS

Thanassoulis and Dyson (1992) developed models that can be used to estimate alternative input-output target levels to render relatively inefficient DMUs efficient. These models can also incorporate preferential weights on the input-output improvements such that the final target levels reflect the user’s preferences in achieving efficiency through alternative paths. Then the authors discuss the estimation of targets when one of the input or output is given a pre-emptive priority to improve. As a generalization of the above, the authors also discuss a model with a general preferential structure that allows for a different degree of importance to be attached to the improvements in each of the input and outputs. Further, the model also accounts for the improvements in the inputs and the outputs simultaneously. An important aspect in the formulation of the model is that, the structure does not necessarily require proportional changes in the case of all inputs or all outputs, hence making it different from the classical radial models. Different factors are used in the constraints for the increase of each of the outputs and the decrease of each of the inputs. These factors are then weighted appropriately and then included in the objective function. This model with the preferential weights on the improvements of the inputs and the outputs is shown in the full form in the linear program in 2.45

$$
\begin{aligned}
\max \quad & \sum_{r\in R_0} w_r^{+}\boldsymbol{b}_r - \sum_{i\in I_0} w_i^{-}\boldsymbol{a}_i + \boldsymbol{e}\Big(\sum_{i\in \bar{I}_0} d_i^{-} + \sum_{r\in \bar{R}_0} d_r^{+}\Big) \\
s.t. \quad & \boldsymbol{b}_r y_{rj_0} - \sum_{j=1}^{n} z_j y_{rj} = 0, \quad r\in R_0 \\
& \boldsymbol{a}_i x_{ij_0} - \sum_{j=1}^{n} z_j x_{ij} = 0, \quad i\in I_0 \\
& \sum_{j=1}^{n} z_j y_{rj} - d_r^{+} = y_{rj_0}, \quad r\in \bar{R}_0 \\
& \sum_{j=1}^{n} z_j x_{ij} + d_i^{-} = x_{ij_0}, \quad i\in \bar{I}_0 \\
& \boldsymbol{b}_r \geq 1 \quad \forall \quad r\in R_0 \\
& \boldsymbol{a}_i \leq 1 \quad \forall \quad i\in I_0 \\
& z_j \geq 0 \quad \forall \quad j \\
& \boldsymbol{a}_i, \boldsymbol{b}_r \;\; free \,\forall\, i\in I_0 \;\; and \;\; r\in R_0 \\
& d_i^{-}, d_r^{+} \geq 0 \quad \forall \quad i\in \bar{I}_0 \;\; and \;\; r\in \bar{R}_0
\end{aligned}
\qquad \textbf{(2.45)}
$$

where,

n – Number of DMUs assessed

I – Index set of inputs, $I = \{1,2,\ldots m\}$

R – Index set of outputs, $R = \{1,2,\ldots s\}$

x_{ij} – Input *i ÎI, j = 1,2,…n*

y_{rj} – Output *r ÎR, j = 1,2,…n*

b_r – Factor of increase of the outputs

a_I – Proportion of decrease of the inputs

w_r^+ – Preferential weights attached to the factor of increase of the outputs

w_I^- - Preferential weights attached to the proportion of decrease of the inputs

d_I^- - Deviation variable associated with the inputs

d_r^+ - Deviation variable associated with the outputs

e - Small positive number

z_j – Weights associated with DMU j

There might be DMUs that would be able to articulate their targets they wish to adopt. In these cases, each of the DMUs would have ideal targets that they would wish to achieve. These targets might sometimes be in such a way that the DMU would be willing to sacrifice the level of some of the input(s) and/or output(s) in the process of improving the others. Some of these ideal targets in general might be neither feasible nor efficient. Hence the authors propose a two-stage approach to determine feasible and efficient targets.

In the first stage, feasible input-output levels are determined that are as close as possible to the ideal targets. In the second stage, the model is used to check for the presence of a set of efficient input-output levels that dominates the input-output levels in stage 1. This set of input-output levels determined in stage 2 are said to be compatible with the original ideal targets and are used as the targets for the DMU.

The linear program is as in 2.46.

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{m} w_i^{1-} k_i^1 + \sum_{i=1}^{m} w_i^{2-} k_i^2 + \sum_{r=1}^{s} w_r^{1+} c_r^1 + \sum_{r=1}^{s} w_r^{2+} c_r^2 \\
s.t. \quad & \sum_{j=1}^{n} g_j y_{rj} + c_r^1 - c_r^2 = y_r^t \ , \ r = 1,2,\dots s \\
& \sum_{j=1}^{n} g_j x_{ij} + k_i^1 - k_i^2 = x_i^t \ , \ i = 1,2,\dots m \\
& g_j \geq 0 \quad \forall \quad j \\
& c_r^j, k_i^j \geq 0 \quad for \ \ j = 1,2 \quad and \ \forall \ i \ and \ r
\end{aligned}
\tag{2.46}
$$

The formulation in 2.46 is a preferential structure formulation that tends to minimize the positive and the negative slacks (deviational variables) associated with the inputs and the outputs. The values on the right hand side are the targets levels associated with the input and the output variables. Also the deviation from the targets are associated with a penalty which are the *‘w’* s in the objective function. In the next chapter, the goal programming formulation will be used to address the questions that we raised with respect to accounting for the preference for the improvement of the different types of variables like the desirable and the undesirable outputs.

Thanassoulis and Dyson (1992) formulation does not address planning and resource allocation issues at the global organizational level while considering all DMUs simultaneously. Athanassopoulos (1995) provided enhancements to that formulation by including global organizational targets and global resource constraints. However, the model developed by Athanassopoulos (1995) is not represented here, as the enhancements provided do not have any direct relation with the formulation to be discussed in the methodology in Chapter 3. For further information, the reader is directed to the original paper.

Finally, one important aspect to be noted is that none of the models discussed so far accounted for the presence of any sort of technological dependencies between the desirable outputs and the undesirable outputs. In many real life scenarios, in addition to weak disposability and the null jointness between the desirable and the undesirable outputs, there might exist a relationship between the two types of outputs. In the next Chapter, a more detailed explanation of the situation and the methodology of the formulation to account for the relationship will be discussed.

CHAPTER 3. Methodology

In the previous chapter we discussed a few direct and indirect approaches in incorporating undesirable outputs in the model while determining the efficiency of the DMUs. We also consider these outputs differently from the other outputs in such a way that when an increase in the desirable outputs is desired, a simultaneous reduction in the undesirable outputs is modeled. Finally, we discussed the Goal Programming approach that is being used in Data Envelopment Analysis and also introduced the concept of GoDEA.

In this Chapter, we will initially discuss the interdependencies between the desirable and the undesirable outputs and as to how these can be incorporated in the model. Later we will present how goal programming can be used to check and determine the targets for the DMUs.

3.1. Interdependencies between Desirable and Undesirable Outputs

Zofio and Prieto(2001) in their paper talk about the existence of congestion. Then they go about to talk in detail about how congestion binds the production process if the authority sets a regulatory standard. The authors explicitly introduce a constraint to account for the regulatory standard in the linear programming formulation. There is no discussion about the presence of any technological dependence between the desirable and the undesirable outputs.

In the case of certain transformation processes, it is possible that the Decision-maker knows a priori a standard or an achievable relationship with regard to the desirable and the undesirable outputs. The Decision-maker would then set this relationship as a target to be achieved. This is driven by the existence of a technological relationship between the desirable and the undesirable outputs and that he/she can explicitly determine the relationship in the form of a linear dependence. A typical example for this can be found in the transformation process in the chemical industry.

According to a given chemical process, let us assume that along with the required product, there is also a waste product. By increasing the amount of the required product being produced, one will end up with a proportional increase in the waste product as well.

3.1.1. One Desirable output – One Undesirable output

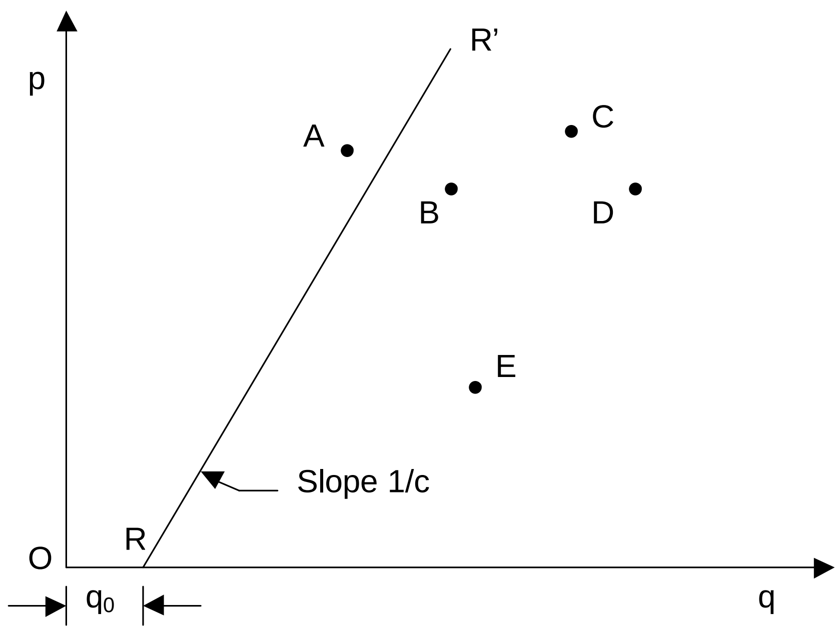


Figure 3.1 Linear Dependence relationship

Let us assume that A through E are the observed DMUs as shown in the Figure 3.1. These DMUs are plotted in a desirable-undesirable output space pOq. Let RR’ be the line representing the technological relationship between the desirable and the undesirable output. Here ‘q_0’ is the residual amount of the undesirable output that would be produced without any production of the desirable output.

A typical example of observing such a residual amount of the undesirable output is the printing of papers in the Pressroom at the Washington Post. At the start of any production run, the presses start with an initial amount of waste papers being printed. This is often due to one or more reasons. One such reason is due to excess or insufficient ink applied onto the plates. This causes the papers that are printed to be blotted with ink in the former case. In the latter case, the text printed on the paper is too light to read. Either of these requires the attention of the Pressroom personnel to attend to the inking process and rectify it. The presses usually are in running state and are not stopped for any rectification of the inking process. Due to not stopping during this time, the presses print some paper that need to be discarded. Only after the rectification, good papers start to be printed. Hence we see that initially there is a residue of waste when the number of good papers produced is zero.

'*1/c*' is the slope that represents the gradient of the linear relationship between the desirable and the undesirable outputs after the initial residual amount. Here '*c*' represents the amount of undesirable output that can be produced for every unit of the desirable output. Again in the case of the Washington Post Pressroom example, the Quality Control person will have a rate of $c = 0.003$. This means that for every 1000 good papers printed, 3 bad papers can be printed.

The decision maker would expect all the DMUs to lie to the north-west of the line RR' (like DMU A), meaning that each of the DMU need to be producing equal or greater amount of the desirable output than that determined by the line RR' for a given amount of undesirable output. Or in other words, each DMU should produce equal or lesser amount of the undesirable output than that determined by the line RR' for a given amount of desirable output. Hence typically the requirement would have the following form,

$$q \leq cp + q_0 \tag{3.1}$$

where,

q – *Amount of undesirable output*

$1/c$ – *Slope of the line depicting the relationship*

p – *Amount of desirable output*

q_0 – *Residual amount of undesirable output*

But in reality some of the DMUs that are not operating as required would end up lying on the south-east side of the line RR' such as DMUs B, C, D and E in Figure 3.1. The decision-maker would also be interested in determining the efficiency score in the presence of the interdependency and the extent to which this relation binds the production process. To take into account these issues, the previously defined technological relationship needs to be incorporated into the model in such a way that targets that are set for the non-efficient DMUs need to satisfy the technological dependence requirement. This can be done by determining hypothetical DMUs that lie to the northwest of the line RR' for the purpose of comparison.

By this procedure of incorporating the new constraint, one ensures that whenever the efficiency of a DMU is evaluated, it is compared to DMUs (real or hypothetical) on or to the northwest of the line RR'. This step might increase the reference set (peers) of the DMUs. This is because, by adding the new constraint, we will be comparing each DMU with a hypothetical or an observed DMU that is efficient and satisfies the constraint as well.

Introduction of this technological dependence relation into the model would bind the production process and one can complete a sensitivity analysis as to whether the evaluation scores change dramatically. This is possible by the comparison of the efficiency scores before and after the introduction of the relation.

3.1.2. Nomenclature

In the previous section we discussed in detail the dependence of a single desirable output and single undesirable output. We also looked at the intuition behind the inclusion of these relations into the model. Here, let us start with the required nomenclature for the general case and the model formulations explained later in this Chapter.

Let us assume that a set of DMUs $j=\{1,\ldots J\}$ use a set of x inputs to produce y outputs out of which p are desirable and q are undesirable.

$$
\begin{aligned}
x &= (x_1, x_2, \dots x_M) \in \Re_+^M \\
p &= (p_1, p_2, \dots p_N) \in \Re_+^N \\
q &= (q_1, q_2, \dots q_R) \in \Re_+^R \\
y &= (p, q) \in \Re_+^{N+R}
\end{aligned}
\qquad \textbf{(3.2)}
$$

The subsets *p* and *q* of set *y* are mutually exclusive and collectively exhaustive of the set *y*.

The technology set *T* consists of all vectors *(x, y)* i.e.,

$$T = \{(x, y) : x \; can \; produce \; y = (p, q)\} \qquad \textbf{(3.3)}$$

Assumptions on the set *T*:

- The set *T* is a closed set.
- The set of inputs *x* and the set of desirable outputs *p* are strongly or freely disposable.
- Weak disposability of the undesirable outputs is assumed.

  If *(x, y) Î T* and *0 £ d £ 1* then *(x,dy) Î T*
- Null-jointness

  If *(x, y) Î T* where *y = (p, q)* and *q = 0* then *p = 0*

3.1.3. Dependence in a multiple outputs scenario

Using the equation in 3.1, the single desirable single undesirable output case for a set of *j=1,2,…J* DMUs, the relationship is

$$\sum_{j=1}^{J} \boldsymbol{a}_j q_j \le c \sum_{j=1}^{J} \boldsymbol{a}_j p_j + q^0 \qquad \textbf{(3.4)}$$

When the value of ‘c’ is forced to zero by the decision-maker, then the above inequality reduces to

$$\sum_{j=1}^{J} \alpha_j q_j \leq q^0 \qquad \textbf{(3.5)}$$

which is the equivalent of ‘Qz £ q_{rk}’ as given by Zofio and Prieto (2001)

Here we take a step further to account for the relationships between each of the undesirable outputs with each and every desirable output.

$$\begin{aligned}
&\sum_{j=1}^{J} \alpha_j q_j \leq c'_{11} \sum_{j=1}^{J} \alpha_j p_{1j} + q_1^0 \\
&\sum_{j=1}^{J} \alpha_j q_j \leq c'_{12} \sum_{j=1}^{J} \alpha_j p_{2j} + q_2^0 \\
&. \\
&. \\
&. \\
&\sum_{j=1}^{J} \alpha_j q_j \leq c'_{1n} \sum_{j=1}^{J} \alpha_j p_{nj} + q_n^0
\end{aligned} \qquad \textbf{(3.6)}$$

Summing up all the inequalities in 3.6,

$$n\sum_{j=1}^{J} \alpha_j q_j \leq c'_{11} \sum_{j=1}^{J} \alpha_j p_{1j} + c'_{12} \sum_{j=1}^{J} \alpha_j p_{2j} + \ldots c'_{1n} \sum_{j=1}^{J} \alpha_j p_{nj} + q_1^0 + q_2^0 + \ldots q_n^0 \qquad \textbf{(3.7)}$$

The inequality 3.7 can be reduced further and in its generalized sense (n – desirable and r – undesirable outputs) takes the form as below in the inequality 3.8

$$\sum_{j=1}^{J} \alpha_j q_{rj} \leq \sum_{n=1}^{N} c_{rn} \sum_{j=1}^{J} \alpha_j p_{nj} + q_r^0 \quad \forall \quad r = 1,2,\ldots R \qquad \textbf{(3.8)}$$

The relationship as such is given by the parameters c_{rn} and q_r^0. Since the parameters are known, these are applied into the constraint. The values of c represent the gradients of each of the desirable outputs in relation to each of the undesirable outputs. Hence it is a matrix of dimension *rxn*. The intercepts on each of the undesirable output axes are given by values of q^0. It is a matrix of dimension *r*.

The inequality in 3.8 thus puts forth the constraint that the hypothetical DMU used for the comparison, cannot have an amount of the undesirable output that exceeds the amount determined by the expression on the right hand side of the inequality. The value of this expression on the right hand side is in turn determined by the values of the parameters. This constraint is applied to each one of the undesirable outputs.

The inequality developed in this section can be incorporated into models that are used for efficiency measurement. In the next few sections, we will modify the Index number approach developed by Fare et al. (2000) and the Hyperbolic approach originally introduced by Fare et al. (1985) and recently modified by Zofio and Prieto(2001).

3.1.4. Interdependency Incorporated in the Hyperbolic Efficiency Measurement Model

The modified hyperbolic approach for the measurement of efficiency by Zofio and Prieto (2001) as explained in Chapter 2 looks at the expansion desirable outputs by a factor q and a contraction in the undesirable outputs by a factor q.

Here we apply the inequality that we developed in 3.8 as a constraint to the model with weak disposability assumption only. It can similarly be incorporated in the model with the strong disposability assumption.

The formulation is as in 3.9 below.

Model A:

$$
\begin{aligned}
&\max \quad q' \\
&s.t. \\
&\sum_{j=1}^{J} a_j p_{nj} \geq q' p_{nj_0} \quad \forall \quad n = 1,2,\ldots N \\
&\sum_{j=1}^{J} a_j q_{rj} = q_{rj_0} / q' \quad \forall \quad r = 1,2,\ldots R \\
&\sum_{j=1}^{J} a_j x_{mj} \leq x_{mj_0} \quad \forall \quad m = 1,2,\ldots M \\
&\sum_{j=1}^{J} a_j q_{rj} \leq \sum_{n=1}^{N} c_{rn} \sum_{j=1}^{J} a_j p_{nj} + q_r^0 \quad \forall \quad r = 1,2,\ldots R \\
&j_0 = 1,2,\ldots J \\
&q', a_j \geq 0 \quad \forall \; j = 1,2,\ldots J
\end{aligned}
\tag{3.9}
$$

The linear program 3.9 is in the envelopment form, where together with the scores of the individual DMUs, the weights of the peers are also determined. As explained before in Chapter 2, the weights of the peers determine the extent to which each of the other DMUs are important in coming up with a hypothetical DMU on the efficient frontier for the purpose of comparison with the DMU being evaluated. It can be seen that, q' is the factor of increase of the desirable outputs and the decrease of undesirable outputs. The objective function looks at maximizing the factor q'.

The first three constraints are similar to the constraints from the original formulation by Zofio and Prieto(2001) (refer to Chapter 2). The set of a_j are the weights of the peers. The remaining variables are as mentioned in the nomenclature.

The last constraint is effectively the difference between the original model by Zofio and Prieto(2001) and the model described here. The inclusion of the additional constraint accounts for the existence of a previously known relationship between the desirable and the undesirable outputs. This relationship as explained before is an ideal and expected linear relationship. The effect of the ideal performance relationship can be studied by computing the ratio, $q¢ /q^{Z}$ where, q' is the factor of increase of the desirable and the decrease of the undesirable outputs.

3.1.5. Interdependency incorporated in the Index number approach

The index number approach developed by Fare et al. (2000) as explained earlier is a two step approach to determine the efficiency scores. Each step computes an index number. The first step determines the quantity index of desirable outputs and the second, the quantity index of undesirable outputs. The overall environmental performance index is given by the ratio of these two indices.

The inequality constraint can also be incorporated into this model as below in 3.10

<u>Model B</u>:

Desirable outputs:

$$
\begin{aligned}
&\max \quad \boldsymbol{q} \\
&s.t. \\
&\sum_{j=1}^{J} \boldsymbol{a}_j p_{nj} \geq \boldsymbol{q}\, p_{nj_0} \quad \forall \quad n = 1,2,\ldots N \\
&\sum_{j=1}^{J} \boldsymbol{a}_j q_{rj} = q_{rj_0} \quad \forall \quad r = 1,2,\ldots R \\
&\sum_{j=1}^{J} \boldsymbol{a}_j x_{mj} \leq x_{mj_0} \quad \forall \quad m = 1,2,\ldots M \\
&\sum_{j=1}^{J} \boldsymbol{a}_j q_{rj} \leq \sum_{n=1}^{N} c_{rn} \sum_{j=1}^{J} \boldsymbol{a}_j p_{nj} + q_r^0 \quad \forall \quad r = 1,2,\ldots R \\
&\boldsymbol{q}, \boldsymbol{a}_j \geq 0 \quad \forall\ j = 1,2,\ldots J \\
&j_0 = 1,2,\ldots J
\end{aligned}
\qquad \textbf{(3.10)}
$$

Undesirable outputs:

$$
\begin{aligned}
&\min \quad \lambda \\
&s.t. \\
&\sum_{j=1}^{J} \alpha_j p_{nj} \geq p_{nj_0} \quad \forall \quad n = 1,2,\ldots N \\
&\sum_{j=1}^{J} \alpha_j q_{rj} = \lambda q_{rj_0} \quad \forall \quad r = 1,2,\ldots R \\
&\sum_{j=1}^{J} \alpha_j x_{mj} \leq x_{mj_0} \quad \forall \quad m = 1,2,\ldots M \\
&\sum_{j=1}^{J} \alpha_j q_{rj} \leq \sum_{n=1}^{N} c_{rn} \sum_{j=1}^{J} \alpha_j p_{nj} + q_r^0 \quad \forall \quad r = 1,2,\ldots R \\
&\theta, \alpha_j \geq 0 \quad \forall\ j = 1,2,\ldots J \\
&j_0 = 1,2,\ldots J
\end{aligned}
\qquad \textbf{(3.11)}
$$

The assumptions for the above two models 3.10 and 3.11 are weak disposability of the undesirable outputs; i.e. the undesirable outputs can be reduced only when accompanied by a proportional reduction in the desirable outputs.

The two models have the additional constraint incorporated that modifies the computed score of the DMUs accounting for their performance in the presence of the ideal relationship. The desirable output and the undesirable output quantity indices are computed as in Chapter 2 with respect to a specific reference DMU and then the overall environmental performance index is given as the ratio of the desirable output quantity index to the undesirable output quantity index.

So far we discussed the computation of the efficiency scores in the presence of interdependencies of the desirable and the undesirable outputs that we have known a priori. In the next section we will look at a modified approach that helps the decision-maker to determine the actual interdependency.

3.2. Determination of the interdependencies

The original formulation by Charnes et al. (1978) was developed to determine the efficiency score of the transformation process associated with the DMUs. It evaluated the performance of the DMUs by computing scores that were based on the amount of resources used and the amount of outputs produced.

Here we deviate from the basic intuition of the transformation process and come up with a different type efficiency score based on the relationship between the desirable and the undesirable outputs. The intuition arises with the basic assumption that a linear relationship exists between the two types of outputs. Using this assumption, a different type of efficiency score r is defined as below.

$$r = \frac{\text{Observed amount of desirable output}}{\text{Ideal amount of desirable output}} \qquad \textbf{(3.12)}$$

The next question that needs to be answered is a way to compute the ideal amount of the desirable output. This can be done by going back to the linearity assumption used to determine the ideal amount of desirable output using the observed amount of undesirable output. It can be given as in 3.13 below.

$$\text{Ideal amount of desirable output} = aq + p^0$$

where,

a – slope of the gradient of the relationship between the desirable and the undesirable output

p^0 – the intercept with the desirable output axis

Here, a and p^0 are the variables that need to be determined. These are parameters and a parallel could be drawn to the weights for the inputs and the outputs in the original CCR model by Charnes et al. (1978). Then based on the definition, the efficiency score can be given as

$$\boldsymbol{r} = \frac{p}{aq + p^0} \tag{3.14}$$

The above fraction can take any value between 0 and 1. When the value is one, the DMU is environmentally efficient. As the observed amount of the undesirable output *p* decreases from the ideal value, the score decreases from one.

$$p \to 0 \Rightarrow \boldsymbol{r} \to 0 \tag{3.15}$$

Using the expression in 3.14, a linear program can be formulated to compute the efficiency score of each DMU considering the desirable and the undesirable outputs. But, definitely, the efficiency score thus computed would give us insight on how the DMU is performing with respect to the desirable and the undesirable outputs alone. This does not account for the input performance of the DMU in relation to other DMUs.

Typically, assuming the presence of *J* DMUs, the parameters will be determined using the linear program and the DMU will be given the freedom to assign values to parameters that will give a score as close to one as possible. The constraint on the parameters will be that if the parameters are used to compute an equivalent score for any of the DMUs, the score should not fall below one. Hence the program is formulated as below in 3.16.

$$\begin{aligned} &\min \frac{aq_{j_0} + p^0}{p_{j_0}} \\ &st \\ &\frac{aq_j + p^0}{p_j} \geq 1 \quad \forall \; j = 1,2,\dots J \\ &a, p^0 \; unrestricted \end{aligned} \tag{3.16}$$

In the presence of *N* desirable outputs and *R* undesirable outputs, the linear program is modified and the takes the form as in 3.17.

Model C:

$$
\begin{aligned}
&\min \sum_{n=1}^{N} u_n \frac{\sum_{r=1}^{R} a_{nr} q_{rj_0} + p_n^0}{p_{nj_0}} \\
&st \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textbf{(3.17)} \\
&\sum_{n=1}^{N} u_n \frac{\sum_{r=1}^{R} a_{nr} q_{rj} + p_n^0}{p_{nj}} \geq 1 \quad \forall\ j = 1,2,\ldots J \\
&a_{nr},\ p_n^0\ \textit{unrestricted}
\end{aligned}
$$

The above-formulated model gives an efficiency score based on the performance of the DMUs as explained before. Also by solving the above model, one can get values for the parameters for each of the DMUs. These parameters are determined in such a way so as to maximize the performance of each DMU. One will end up with J sets of parameter values from the model. This would give insight into the technological interdependence that exists between the desirable and the undesirable outputs. Based on this, the decision-maker will be able to set new targets for the DMUs in future. Further, these can be used as a correction measure to modify or alter earlier targets that are set based on technological interdependence. It also gives best possible relationships between the various desirable and undesirable outputs in the form of hyperplanes.

So far the models developed and discussed have been non-goal programming approaches to incorporate the effects of undesirable outputs. The last model formulated looked at the performance based on only the desirable and the undesirable outputs. We have not addressed the problem when considering the inputs. In the next section, we will look at models that include the inputs in the measurement process using a preferential weight structure.

3.3. Modified Approach to Incorporate Desirable and Undesirable Outputs with a Preferential Structure

As explained in Chapter 2, Thanassoulis and Dyson (1992) formulated a weight-based preferential structure model that would allow a differential degree of importance to be attached to the potential changes of the individual input or output levels. In this section we develop a modified model based on the preferential weight structure that will incorporate the increase of the desirable outputs and the decrease of the undesirable outputs with preferential weights attached to them.

The model is as shown in 3.18

Model D:

$$
\begin{aligned}
&\max \quad \sum_{n=1}^{N} w_n^p \boldsymbol{s}_n - \sum_{r=1}^{R} w_r^q \boldsymbol{w}_r \\
&s.t. \\
&\sum_{j=1}^{J} \boldsymbol{a}_j p_{nj} = \boldsymbol{s}_n \, p_{nj_0} \quad \forall \quad n = 1,2,\ldots N \\
&\sum_{j=1}^{J} \boldsymbol{a}_j q_{rj} = \boldsymbol{w}_r q_{rj_0} \quad \forall \quad r = 1,2,\ldots R \\
&\sum_{j=1}^{J} \boldsymbol{a}_j x_{mj} \leq x_{mj_0} \quad \forall \quad m = 1,2,\ldots M \\
&\boldsymbol{s}_n \geq 1 \quad \forall \; n = 1,2,\ldots N \\
&0 \leq \boldsymbol{w}_r \leq 1 \quad \forall \; r = 1,2,\ldots R \\
&\boldsymbol{a}_j \geq 0 \quad \forall \; j = 1,2,...J
\end{aligned}
\qquad \textbf{(3.18)}
$$

Unlike the hyperbolic approach initially developed by Fare et al. (1989) and modified by Zofio and Prieto (2001), the basic structure of the model developed in 3.18 allows varying amounts of increase of the desirable outputs and decrease of the undesirable outputs. Compared to the index number approach by Fare (2000), here both the increase of the desirable outputs and the decrease of the undesirable outputs are accounted for in the same model. Also it can be seen that the model in 3.18 has s and w in N and R dimensions respectively. This means that the model has a non-radial structure. The model does not necessarily require increase (decrease) with respect to the desirable (undesirable) outputs to be in the same proportion. Hence the above model could be used to set targets in the context of a generalized preferential structure over potential changes to desirable-undesirable output changes.

s_n is the proportional increase in the n^{th} desirable output and w_r is the proportional decrease in the r^{th} undesirable output. A DMU will be efficient if,

$$s_n^* = w_r^* = 1 \ \forall \ n \text{ and } r$$

The above approach does not look at the changes in the inputs as a contribution to the overall performance. Hence the model needs to be modified in a way to account for the performance of the DMUs with respect to all the three types of variables- inputs, desirable outputs and undesirable outputs. In the next section, we look at a model that addresses this issue.

The values of the weights w in the objective function are set by the decision-maker based on experience and the requirements. The values of these represent the relative importance of one output (desirable or undesirable) with respect to the rest of the variables.

3.4. A Goal Programming approach

The Goal Programming approach in DEA is a method formulated by Athanassopoulos (1995) to account for issues in the Thanassoulis and Dyson (1992) model explained in Chapter 2. The model explained in this section minimizes the slacks associated with the inputs, desirable and undesirable outputs.

Model E:

$$
\begin{aligned}
&\min \quad \sum_{n=1}^{N}(w_n^{p+}s_n^{p+} + w_n^{p-}s_n^{p-}) + \sum_{r=1}^{R}(w_r^{q+}s_r^{q+} + w_r^{q-}s_r^{q-}) + \sum_{m=1}^{M}(w_m^{x+}s_m^{x+} + w_m^{x-}s_m^{x-}) \\
&s.t. \\
&\sum_{j=1}^{J}\boldsymbol{a}_j p_{nj} + s_n^{p+} - s_n^{p-} = p_{nj_0} \quad \forall \quad n = 1,2,\ldots N \\
&\sum_{j=1}^{J}\boldsymbol{a}_j q_{rj} + s_r^{q+} - s_r^{q-} = q_{rj_0} \quad \forall \quad r = 1,2,\ldots R \\
&\sum_{j=1}^{J}\boldsymbol{a}_j x_{mj} + s_m^{x+} - s_m^{x-} = x_{mj_0} \quad \forall \quad m = 1,2,\ldots M \\
&s_n^{p+}, s_n^{p-} \geq 0 \quad \forall\, n \\
&s_r^{q+}, s_r^{q-} \geq 0 \quad \forall\, r \\
&s_m^{x+}, s_m^{x-} \geq 0 \quad \forall\, m \\
&\boldsymbol{a}_j \geq 0 \quad \forall\; j = 1,2,\ldots J
\end{aligned}
\tag{3.19}
$$

The model in 3.19 has a Goal programming structure because of the structure of the objective function, where the overall target is to minimize any inefficiency associated with any of the variables. This is a modification of the model explained in Chapter 2. The model gets to the frontier by minimizing the positive and negative deviational variables associated with the inputs, desirable and the undesirable outputs. The *s* are the deviational variables. Each input or desirable output or undesirable output has one positive and one negative deviational variable associated with it.

This model is different from the rest of the models developed in this Chapter because, unlike other models, this model looks at improving the performance by considering the inefficiency with respect to all the three different types of variables. The *w* are the weights associated with the deviational variables to account for relative importance in the improvement of the performance respective variables. The values of these weights again are determined by repeated solving of the model and by experience.

CHAPTER 4. Application, Results, And Discussion

In this chapter, we illustrate the four formulations developed in Chapter 3 for the treatment of undesirable outputs. We do this by applying them to three data sets, two from the DEA literature and another one collected with the purpose of this research. The results obtained from other models will be used to compare the performance of the formulations developed here. Table 4.1, at a glance, provides information about the source of the data set used for illustrating each model.

Whenever possible, the models developed were applied to the same data sets used in the literature to illustrate the basic formulations that we have modified. This is the case for models A and B, which were compared to their corresponding models in the literature — models L1 and L2. As model C presents a formulation for the determination of interdependencies that has not been attempted, the results from this formulation was not compared to any other. Model D was compared to Fare's index number approach model, as both have individual measures for capturing inefficiency relating to the desirable and undesirable outputs. Model E is compared to models L1 and L2 since both formulations give also a measure of the slacks.

Model	**Source of Data Set**	**Model to be compared to**
Model A: Interdependency incorporated in the Hyperbolic Efficiency Measurement Model (see section 3.1.4)	*Data from the manufacturing industries of 14 OECD countries*	Model L1: Hyperbolic Efficiency Measurement Model —Zofio and Prieto (2001)
Model B: Interdependency incorporated in the Index number approach (see section 3.1.5)	*Data from the manufacturing industries of 14 OECD countries*	Model L2: Index number approach —Fare et al. (2000)
Model C: Determination of interdependencies (see section 3.2)	*Undesirable Outputs in Efficiency Valuations* —Scheel (2001) *Data from the manufacturing industries of 14 OECD countries*	
Model D: Modified approach to incorporate desirable and undesirable outputs with a preferential structure (see section 3.3)	*Data from the manufacturing industries of 14 OECD countries*	Model L2: Index number approach —Fare et al. (2000)
Model E: Goal Programming Approach to incorporate undesirable outputs (see section 3.4)	*Data from the manufacturing industries of 14 OECD countries* *Newspaper print production data 2001*	Model L1: Hyperbolic Efficiency Measurement Model —Zofio and Prieto (2001)

Table 4.1 **- Sources of Data Sets for Illustrating the Models for the Treatment of Undesirable Outputs.**

This chapter is divided into two sections. The first section discusses the available data sets. The second section presents the application of each of the five types of models to the data sets described. Finally, this section is concluded by comparing the results of each model with those of the corresponding formulations found in the literature.

4.1. Data Sets

Three data sets are used to compare the performance of the developed formulations against published results from other models. Comparisons against results obtained from applying other models in the literature to these data sets are also performed. The first two data sets describe production from a macro-economic standpoint, since they describe countries. The third data set gives a micro-economic description of a production process. The different character of the data sets illustrates the applicability of the models at different levels and for different purposes.

A short description and background information of each data set follows.

4.1.1. Data from the manufacturing industries of 14 OECD countries

Zofío and Prieto (2001) analyzed the effects of regulatory standards on production, and the limits beyond which production is impossible for different regulatory scenarios on CO_2 emissions. They used data from the Organization for Economic Cooperation and Development (OECD)'s manufacturing industries. The authors did not publish the data but provided the data set for our use in the course of the present research.

This data set contains two inputs, one desirable output and one undesirable output for 14 countries. Table 4.2 presents these variables and some descriptive statistics. The actual data tables are included in Appendix I.

Variable	Type of variable	Unit	Mean	Standard Deviation	Min	Max
Production	Desirable Output	Millions of US$	558555.67	797443.81	32236.07	2861273.00
CO_2 emissions	Undesirable Output	Millions of Tons	133.38	228.43	8.61	886.25
Capital Stock	Input	Millions of US$	248274.47	320755.45	22015.83	1152700.00
Labor	Input	Thousands	3969.66	4997.76	271.80	17496.00

Table 4.2 **- Variables and Descriptive Statistics manufacturing industries of 14 OECD countries**

4.1.2. Undesirable Outputs in Efficiency Valuations —Scheel (2001)

Scheel (2001) applied the efficiency concept to selected European economies adopted from Statistisches Bundesamt (1997). The data set has one input, one undesirable output and one desirable output for 13 countries as shown in Table 4.3. This author acknowledges a high correlation between the desirable and undesirable output of 0.95.

Variable	Type of variable	Unit	Mean	Standard Deviation	Min	Max
Gross Domestic product (GDP)	Desirable Output	DM	901.15	1023.00	92.20	3457.40
NO_x emissions	Undesirable Output	Millions of Tons	174.92	186.00	24.00	612.00
Employees	Input	Thousands	11120.85	11173.49	1262.00	35782.00

Table 4.3 **- Variables and Descriptive Statistics —Scheel (2001)**

4.1.3. Newspaper print production data 2001

The Pressroom at The Washington Post will be used as the research site for this work. Data is being collected to analyze the performance over time of this production department.

The Washington Post is the second newspaper in the U.S. in circulation. Each week the Post Circulation Department oversees delivery of over five million newspapers to homes, businesses and newsstands.

Unit of Analysis:

The concentration will be only on the "Late Run" production.

There are two possible units of analysis or Decision-Making Units to be studied. One is the weeks in a year (analyzing the production process on a weekly basis). For this purpose, the data that is collected daily is aggregated over a week. This is done to remove any non-homogeneity due to the differences in the production run on different days of the week. The data can also be aggregated to a monthly basis in case there is a need to analyze monthly production against weekly production.

A second possible category of DMUs is the daily production of the presses themselves (a total of four presses) with data of their performance over several days.

The latter is done in this research. Figure 3.5 shows the input-output diagram for these processes.

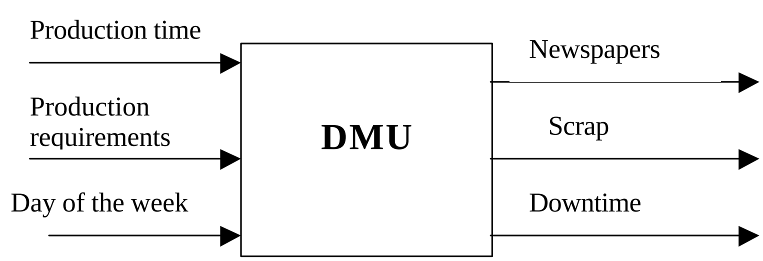


Figure 4.1 Input-Output analysis for newspaper Pressroom

Data Collection for Inputs and Outputs:

Data for inputs and outputs has been obtained from the Press Department at the Washington Post. The Pressroom reports these measures on a regular basis. Table 4.4 presents the available inputs, desirable and undesirable outputs production data. This is data that is being collected for a simulation project aimed at identifying bottlenecks and correcting their causes.

Labor data will not be included in the model because in this application, this variable is fixed for all runs across all days. Production teams are always the same. Capital data is not available at this level. Material data might be possible to collect, and the possible inclusion of the amount of materials input to production each day might be considered as an additional input. At the moment, though, this data is still not available.

The need to ensure that the measures used in the DEA model reflect the relevant perspectives of the organization has been emphasized in the literature (Rouse et al., 1997). The validity of DEA results is often supported by testimony of their usefulness by managers within the organization. Validation by managers of the variables selected for analysis is being sought.

The data set corresponds to the production of four presses on Thursday late night runs for the months of January and February 2001. In total, there are 32 DMUs. The data set contains one input, one desirable output and two undesirable outputs. A description of these variables follows in Table 4.4.

Variable	Type of variable	Unit	Mean	Standard Deviation	Min	Max
Good papers	Desirable Output	Papers	103035.84	22948.88	30981.00	143875.00
Waste	Undesirable Output	Papers	14445.00	759.14	660.00	4600.00
Press downtime	Undesirable Output	Minutes	35.81	23.90	0.00	119.00
Pressrun Time	Input	Minutes	147.50	27.88	65.00	205.00

Table 4.4 **- Variables and Descriptive Statistics —Newspaper pressroom data**

4.2. Programming and Related Issues

4.2.1. Software used

The software used for the programming and the running of the models is Excel Standard Solver and Excel Premium Solver Plus V3.5 as an add-in to Excel 97 on a Windows 2000 platform. The codes and the formulation of the linear programs are illustrated in Appendix II of this document.

4.2.2. Scaling Issues in the data sets

The data sets described above in section 4.1 have variables that vary in magnitude. This gives rise to scaling issues which is taken care of by re-scaling. The software used, causes the coefficients to be re-scaled across both rows and columns, so that all matrix elements are of a similar scale. The scaling problem is very unlikely to occur with Excel 97 or with the enhanced Solver product, especially because of the use of a scaling option. These Solvers include both the ability to automatically re-scale linear models internally, and a more robust test for linearity, part of which is performed at the beginning of the solution process, rather than at the end by earlier versions of excel and the standard solvers that are packaged with Microsoft Excel. We do not lose any information by re-scaling.

In section 4.2 the models developed and existing models in the literature are applied to the data sets here described.

4.3. Application Results

The following sections present the results obtained for the developed formulations and present statistical comparison of these results against formulations existent in the literature as explained in Table 4.1.

4.3.1. Interdependency incorporated in the Hyperbolic Efficiency Measurement Model

The formulation to model interdependency in the hyperbolic efficiency measurement model (Model A) was applied to two sets of data and compared to Zofío and Prieto's hyperbolic efficiency measurement model (Model L1).

The results for each one of the data sets are shown in the Table 4.5 and 4.6 below.

Results from Data from the Manufacturing Industries of 14 OECD Countries:

Model A and Model L1 were applied to Zofío and Prieto (2001) data set. The results as well as the statistical analysis performed in both results are explained here.

DMU	Efficiency score from model A	Efficiency Score from model L1	Effect of Interdependency Eff.(A)/Eff.(L1)
CAN	1.0000	1.0000	1.0000
USA	I.S.	1.0000	I.S.
JAP	I.S.	0.9565	I.S.
AUS	0.8331	0.8331	1.0000
BEL	0.7371	0.7338	1.0045
FIN	0.6662	0.6285	1.0600
FRA	0.8948	0.8141	1.0991
GER	0.9183	0.9127	1.0061
GRE	0.6368	0.5894	1.0804
ITA	0.9497	0.8781	1.0815
NOR	0.7785	0.7252	1.0735
SWE	I.S.	1.0000	I.S.
UK	0.9159	0.7623	1.2015
SPA	0.8362	0.8071	1.0361

I.S. – Infeasible Solution

Table 4.5 **Comparison of efficiency scores from model A and model L1**

	Model A	*Model L1*
Mean	0.833	0.831
Standard Error	0.035	0.036
Median	0.836	0.824
Standard Deviation	0.117	0.135
Sample Variance	0.014	0.018
Range	0.363	0.411
Minimum	0.637	0.589
Maximum	1.000	1.000
Count	11	14

Table 4.6 **Descriptive Statistics for the efficiency scores in Table 4.5**

By applying the interdependency Model A, 9 DMUs experienced an increase in their efficiency score. Two DMUs remained constant and the model did not find a feasible solution for three of them.

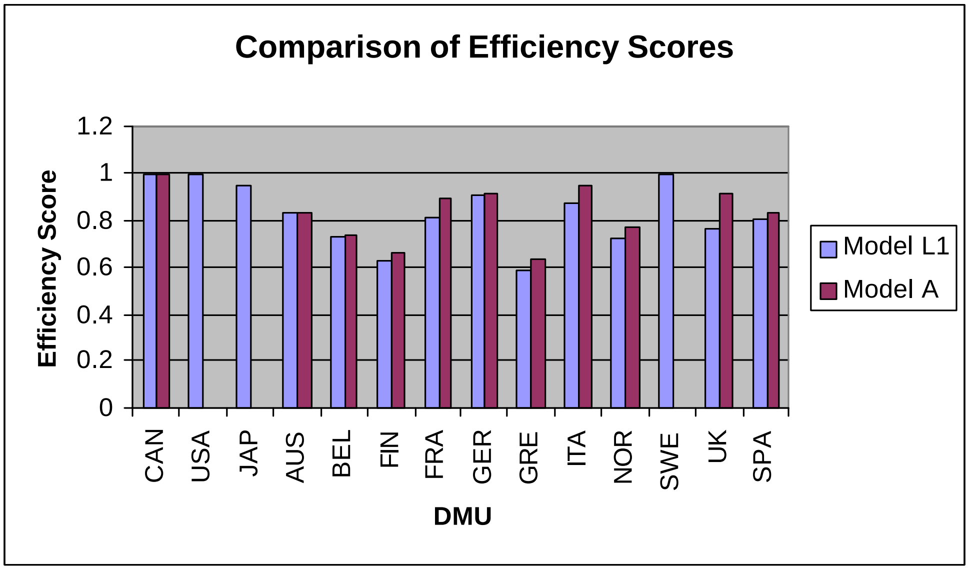


Figure 4.3 Comparison of Efficiency scores from Model L1 and Model A

It is clear that some countries might not be affected by the additional constraint at all, like CAN and AUS. These results are similar to the ones obtained by Zofio and Prieto (2001).

Under the strong disposability assumptions for the model A, for some countries the hyperbolic model cannot reduce the CO_2 emissions by the amount in the restriction because no production would be possible unless at a lower level of efficiency. This can be seen in the case of USA, JAP and SWE where there is no feasible solution.

For the rest of the DMUs, there is an increase in the efficiency score. This might be explained by the fact that Model L1 is already a constrained model, where the search path is restricted to the hyperbolic curve. When adding an additional constraint to model the interdependencies, the solution space gets further reduced. Due to this, either more DMUs fall on the frontier or their path to the frontier gets reduced. This is the justification for the average increase in efficiency scores.

While this feature might seem a disadvantage, it really is an advantage since by incorporating real dependencies between good and bad output we are constraining the solutions to those DMUs that operate under this relationship.

4.3.2. Model B: Interdependency incorporated in the Index number Approach (see section 3.1.5)

The formulation to model interdependency in the index number model (Model B) was applied to Zofío and Prieto's sets of data and compared to Fare's Index number approach (Model L2). The results for each one of the data sets are shown below in tables below.

Results from Data from the manufacturing industries of 14 OECD countries:

Model B and Model L2 were applied to Zofío and Prieto (2001) data set. The results as well as the descriptive statistics in both results are displayed in Tables 4.7 and 4.8 below.

DMU	**Efficiency score from model B**	**Efficiency Score from model L2**	**Effect of Interdependency Eff.(B)/Eff.(L2)**
CAN	0.6171	0.6173	0.9997
USA	0.9686	1.0000	0.9686
JAP	0.9565	1.0000	0.9565
AUS	0.3226	0.3226	1.0000
BEL	0.2730	0.2736	0.9978
FIN	0.2743	0.2880	0.9524
FRA	0.6625	0.7197	0.9205
GER	0.7994	0.8034	0.9950
GRE	0.2839	0.3040	0.9339
ITA	0.7072	0.7581	0.9329
NOR	0.4054	0.4312	0.9402
SWE	1.0000	1.0000	1.0000
UK	0.6740	0.8009	0.8416
SPA	0.6008	0.6203	0.9686

Table 4.7 **Comparison of efficiency scores from model B and model L2**

	Model B	*Model L2*
Mean	0.610	0.639
Standard Error	0.071	0.073
Median	0.640	0.670
Standard Deviation	0.265	0.274
Sample Variance	0.070	0.075
Range	0.727	0.726
Minimum	0.273	0.274
Maximum	1.000	1.000
Count	14	14

Table 4.8 **Descriptive Statistics for the efficiency scores in Table 4.7**

Before comparing the results from the models B and L2, we need to note that model L2 has different assumptions from those of model L1. Further, due to the two-stage evaluation in model L2 the scores are different and cannot be compared to model L1. DMU CAN, which was efficient in model L1, turned out to have an efficiency score of only 0.6173. But at the same time, DMUs USA and SWE remained efficient in both model results.

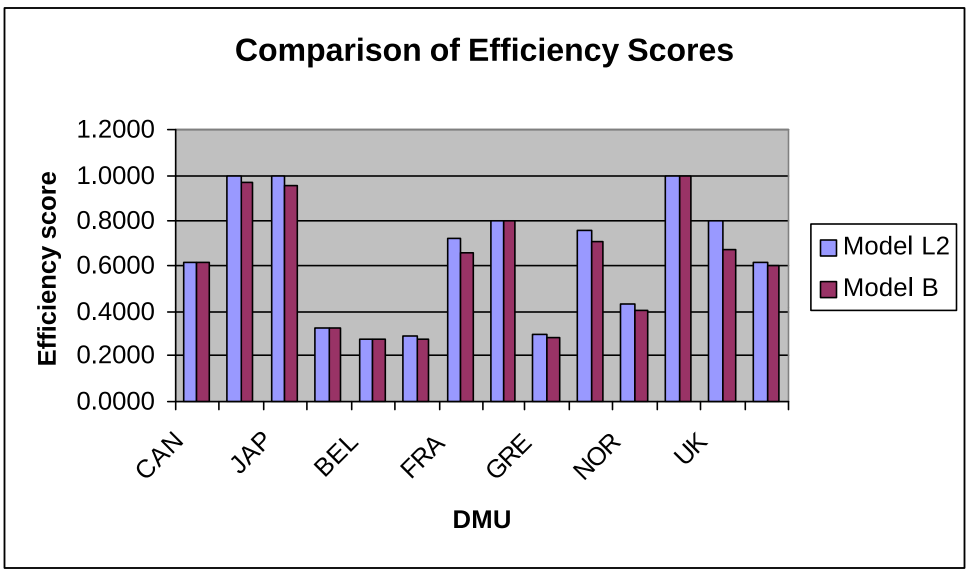


Figure 4. 3 Comparison of Efficiency scores from Model L2 and Model B

In fact, by applying the interdependency Model B, in comparison with model L2, 12 DMUs experienced a decrease in their efficiency score. Two DMUs had no change in their efficiency scores out of which one was inefficient and the other was efficient.

This might be explained because the original model L2 has weak disposability assumption and by incorporating the interdependencies, model B presents strong disposability and the additional technological relationship constraint envelops the data more loosely, thus decreasing the efficiency scores.

When modeling for the interdependencies one can think that units that do not follow the pre-determined relationship will be penalized. The score of the DMU under analysis depends on the added constraint, which determines that the undesirable output of the hypothetical DMU should be less than or equal to the calculated value of the undesirable output. This calculation is done using the interdependency relationship.

4.3.3. Results from Model C: Determination of interdependencies

The formulation to determine interdependencies (Model C) was applied to the data set from Scheel (2001). The results for each one of the data sets are shown in the tables below.

Results from the data in Undesirable Outputs in Efficiency Valuations —Scheel (2001):

Model C was applied to Scheel (2001) data set. The efficiency score obtained from this model evaluates the performance of DMUs considering the production of desirable outputs vs. undesirable outputs, without considering input performance. As such, the results from this model cannot be compared to others for validation purposes. Insights will be obtained from the analysis of a sub-set of the data. For this purpose, the data set has been included in Table 4.9.

Model C produces the equation's parameters —the slope and intercept —of the line that depicts the relationship between desirable and undesirable outputs. Based on those parameters, it is possible to calculate the predicted value of the desirable output. The difference of the predicted to the observed value can provide insights about the validity of model C. The results are in Table 4.10 below.

DMU	GDP	NO_x	Employees	Predicted value of NO_x	% Reduction in NO_x
B	385.7	76	3793	55.7	26.74
D	3457.4	612	35782	461.2	24.64
DK	248.2	61	2601	37.7	38.27
E	802.4	151	12027	110.3	26.96
F	2204.0	294	22057	294.0	0.00
GB	1579.3	456	25936	210.7	53.80
GR	163.8	33	3821	26.6	19.41
I	1560.9	295	19943	208.2	29.42
IRL	92.2	24	1262	17.2	28.29
NL	566.9	116	6782	79.4	31.53
P	144.0	24	4417	24.0	0.00
S	330.9	78	4134	48.5	37.82
SF	179.2	54	2016	28.6	47.01

Table 4.9**. Scheel's data set and predicted value for desirable output**

DMU	Efficiency score from model C
B	0.7133
D	0.7536
DK	0.5822
E	0.7210
F	1.0000
GB	0.4620
GR	0.7702
I	0.7058
IRL	0.6403
NL	0.6702
P	1.0000
S	0.5951
SF	0.4806

Table 4.10 Efficiency Scores from Model C

Interdependency equation:

Slope	Intercept
7.62963	39.11111

Representing the results obtained graphically, we can see the variation in the performance across the different countries in Figure 4.4.

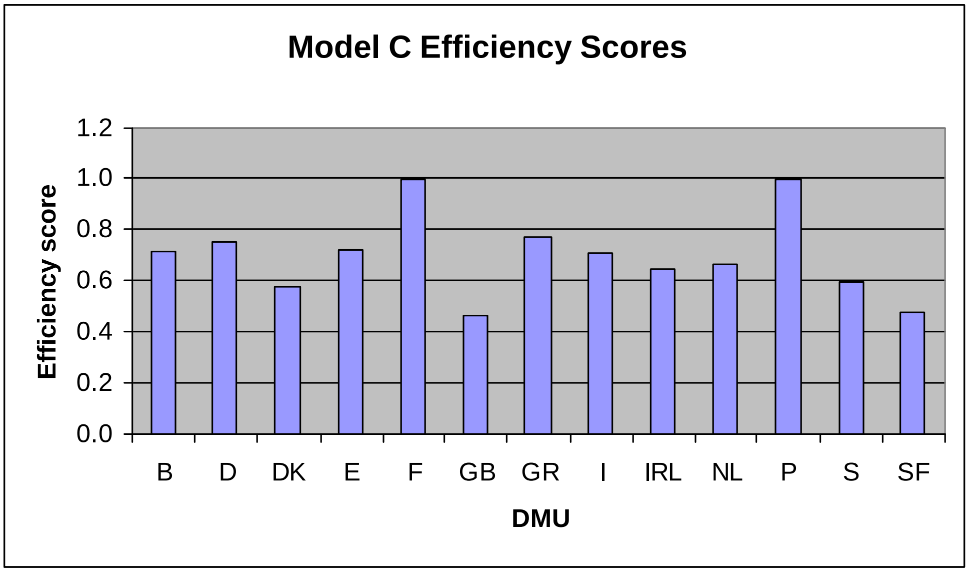


Figure 4. 4 Model C - Efficiency scores

The lowest efficiency score is the one for DMU GB (0.46). A brief observation of the data uncovers that GB is producing 1579.3 units of GDP and 456 units of undesirable output (NOx). For the amount of desirable output that GB is producing it should produce 53.8% less of NO_x to become efficient —which in turn is the highest difference of the data set.

Conversely, the highest efficiency scores are for DMU P and DMU F. Those DMUs produce the exact predicted amount of desirable output relative to the amount of undesirable output.

	Model C
Mean	0.700
Standard Error	0.045
Median	0.706
Mode	1.000
Standard Deviation	0.164
Sample Variance	0.027
Range	0.538
Minimum	0.462
Maximum	1.000
Count	13

Table 4.11 **Descriptive Statistics for the efficiency scores in Table 4.10**

Comparing the results published by Scheel (2001) to the ones obtained, in Scheel's paper DMU P and F are also efficient on an undesirable output oriented non-separating measure. GB is also the lowest efficiency score (0.462). DMUs D and SF are also ranked as 1 with reduced amounts of undesirable outputs. These last results do not compare to Model C since those DMUs are deemed inefficient. This probably can be explained due to the inclusion of inputs in Scheel's formulation. DMU SF, for example, produces more desirable output than DMU GR but using a lot lesser inputs.

4.3.4. Model D: Modified approach to incorporate desirable and undesirable outputs with a preferential structure

The formulation to incorporate desirable and undesirable outputs with a preferential structure (Model D) was applied to Zofío and Prieto's sets of data and compared to Fare's Index number approach (Model L2).

The results for each one of the data sets are shown in tables below.

Results from Data from the manufacturing industries of 14 OECD countries:

Model D and Model L2 were applied to Zofío and Prieto (2001) data set. Both models look at proportional increase of desirable outputs and proportional decrease of undesirable outputs and compute a score based on these two proportions. The scores as well as the statistical description are in Table 4.12 and 4.13 below.

DMU	Efficiency score from Model D	Efficiency Score from Model L2
CAN	0.4340	0.6173
USA	0.5456	1.0000
JAP	0.8775	1.0000
AUS	0.4202	0.3226
BEL	0.4239	0.2736
FIN	0.4940	0.2880
FRA	0.8138	0.7197
GER	0.8758	0.8034
GRE	0.5452	0.3040
ITA	0.7192	0.7581
NOR	0.5552	0.4312
SWE	1.0000	1.0000
UK	0.8842	0.8009
SPA	0.7444	0.6203

Table 4.12 **Comparison in efficiency scores from model D and model L2**

	Model D	*Model L2*
Mean	0.667	0.639
Standard Error	0.054	0.073
Median	0.637	0.670
Standard Deviation	0.200	0.274
Sample Variance	0.040	0.075
Range	0.580	0.726
Minimum	0.420	0.274
Maximum	1.000	1.000
Count	14	14

Table 4.13 **Descriptive Statistics for the efficiency scores in Table 4.12**

By applying the interdependency model D four DMUs experienced a decrease in their efficiency score while nine DMUs experienced an increase. One DMU remained the same (i.e. efficient)

The maximum difference in efficiency scores is for DMU USA, which becomes inefficient to a great extent in model D (0.546). This can be explained by the fact that in model L2, DMU USA dominates while in model D DMU JAP is the peer for USA with a weight of 1.57. Further, we can see that JAP is the only DMU comparable and has lot lesser undesirable outputs (see data in Appendix I and codes for the formulations in Appendix II)

4.3.5. Model E: Goal Programming Approach

The formulation to incorporate desirable and undesirable outputs using goal programming (Model E) was applied to Zofío and Prieto (2001) data set. In the application of the models to the data sets, the weights of the slacks were assumed to be equal to one. This formulation does not provide efficiency scores but identifies the shortfalls and the excesses of the corresponding variables. The results are shown below in tables below.

Results from Data from the manufacturing industries of 14 OECD countries:

Model E was applied to Zofío and Prieto (2001) data set. It is interesting to note that in the goal programming approach, there is no factor to account for the proportional increase or decrease to approach the frontier as in the case of all other models discussed here. Furthermore, all the inefficiency associated to the variables is captured in the slacks. The results (slacks for all variables) were compared to the slacks resulting from the application of the model in Zofío and Prieto (2001) in the table below. The results are as follows:

DMU	Output Shortfall	CO_2 Excess	Capital Stock Excess	Labor Excess
CAN	191.2476	37.01446	0	0
USA	0	0	0	0
JAP	0	0	0	0
AUS	24167.3	12.48727	0	0
BEL	32977	0	0	0
FIN	22551.76	0	1222.782	0
FRA	73189.28	0	20474.84	0
GER	76350.36	0	0	0
GRE	19152.13	0	3251.303	0
ITA	22288	0	138606.5	0
NOR	9272.425	0	4982.642	0
SWE	0	0	0	0
UK	58713.55	0	62614.8	0
SPA	48409.1	0	0	0

Table 4.14 **Results obtained from Model E: Goal Programming Formulation**

DMU	Output Slack	CO2 Slack	Capital Stock Slack	Labor Slack
CAN	0	0	0	0
USA	0	0	3002.94	0
JAP	0	0	14553.58	0
AUS	0	0	0	0
BEL	0	12.30981	0	0
FIN	0	7.348661	1274.532	0
FRA	0	9.949562	25750.21	0
GER	0	9.885347	0	0
GRE	0	5.692634	3286.358	0
ITA	0	2.737925	138888.1	0
NOR	0	2.27793	5011.813	0
SWE	0	0	636.4059	14.356
UK	0	6.802956	67511.14	0
SPA	0	8.480665	516.0885	0

Table 4.15 **Results obtained from model L1: Zofío and Prieto (2001)**

The above comparison is done with the results obtained from Model L1 and not Model L2 because of the fact that Model L2 is a two stage approach and the inefficiencies are appropriately captured in the slacks separately in the two stages. This is unlike Models E and L1, where the slacks are obtained from a single stage.

By applying the goal programming approach of model E the slacks of the desirable output are increased for 11 DMUs while the slacks for CO2 in the goal programming case are zero for 12 DMUs while Model L1 gives values to 9 of them. This is because Model L1 captures part of the inefficiencies in the desirable outputs by way of the factor and the rest through the slacks for each of the variables. On the other hand, Model E captures all of the inefficiencies in the slacks.

The only case where the slacks are comparable is on the input side, where the slacks are very close (see for example DMU FIN that has a value of 1274.5 in Model L1 while in Model E has a value of 1222.7. The same for DMU FRA with value 25750 vs. 20474. Similar values have been obtained for six DMUs in both models while four DMUs had zero slack in both models. Four DMUs had substantially different values.

Further, the slacks for Labor are zero for most DMUs in both models.

Results from Newspaper print production data 2001:

To illustrate the applicability of the goal programming formulation, Model E was applied to the newspaper print production data set. In the application here, the weights for the slacks are assumed to be one. The results are explained in table 4.16 below.

DMU	Output Deficit	Waste Excess	Press Downtime Excess	Press Run Time Excess
1	0	0	0.0	0
2	23303	0	7.6	0
3	27472	0	34.2	0
4	11430	970	12.7	0
5	18390	0	12.8	0
6	0	0	0.0	0
7	26046	0	22.7	0
8	25420	2763	15.1	0
9	29837	0	11.1	0
10	64976	0	35.6	0
11	41777	0	18.8	0
12	18569	0	11.7	0
13	35964	0	17.1	0
14	49232	0	16.1	0
15	0	0	0.0	0
16	26785	0	15.1	0
17	53482	0	54.6	0
18	38061	327	24.2	0
19	10373	0	8.5	0
20	0	0	0.0	0
21	60551	0	38.3	0
22	32935	0	8.0	0
23	51758	0	33.2	0
24	15029	0	9.9	0
25	0	0	0.0	0
26	29340	1914	116.6	0
27	27198	0	15.4	0
28	40016	0	36.9	0
29	28700	0	38.7	0
30	6570	1123	9.3	0
31	25252	0	25.4	0
32	42825	0	38.2	0

Table 4.16 **Results from model E to the newspaper pressroom data**

According to this model, five DMUs do not have any slack in inputs and outputs, thus they can be considered efficient. DMU 26 seem to present the greatest amount of slacks, thus, being the most inefficient. In fact DMU 26 produces less of good papers (30981 units) with a lot of waste (2660 papers) with the maximum downtime of 119 minutes and utilizing 65 minutes of press run time, which is not high. Here the undesirable outputs dominate. The model indicates a slack for the waste as 1914 while 116.6 excess downtime.

DMU 6 has a value of zero for one of the variables — Press Downtime (undesirable output). In the field of DEA, some researchers feel that this might pose a problem as this DMU will dominate the rest due to the high efficiency in that dimension and make all the other DMUs seem to be inefficient. In the application presented in this research, this does not pose any difficulty in the running of the model. The results show that, DMU 6 turns out to be efficient, as expected. This is due to the fact that this DMU is performing efficiently with respect to all the variables — input, desirable and undesirable. In fact, it does not appear as a peer for any of the DMUs. The peers for all the DMUs are given in Table 4.17 below.

DMU	Peer
1	None
2	1, 20
3	1,15
4	20
5	1, 15
6	None
7	1, 20
8	20
9	1, 20
10	1, 15
11	1, 20
12	1, 20
13	1, 20
14	1, 15
15	None
16	1, 15
17	1, 20
18	1, 20
19	1, 20
20	None
21	1, 20
22	1, 20
23	1, 20
24	1, 20
25	None
26	20
27	1, 15
28	1, 20
29	15, 25
30	20
31	1, 15
32	1, 20

Table 4.17 **Model E Peer set for DMUs**

DMU 25 is efficient and appears in the peer set of only DMU 29. DMUs 4, 8, 26 and 30 have only DMU 20 as their peer. All the other inefficient DMUs have two peers in their peer set.

The Figure 4.4 gives a graphical representation of the Output deficits over the 32 DMUs. In this figure five DMUs have zero Output deficits, namely DMUs 1, 6, 15, 20 and 25. This means that these DMUs are performing efficiently with respect to the desirable output – Good papers. These five DMUs are the same DMUs that were declared efficient after analyzing the slacks from all the variables.

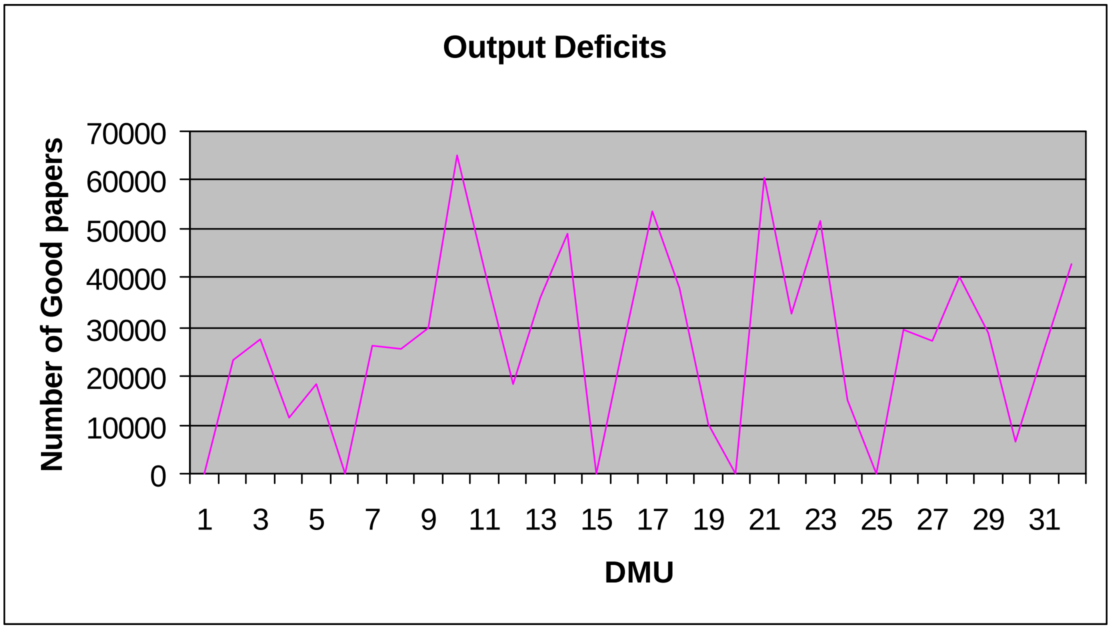


Figure 4. 5 Model E – Output Deficits for the Newspaper Print Production Data

4.4. Implementation of the various Models by the decision-maker

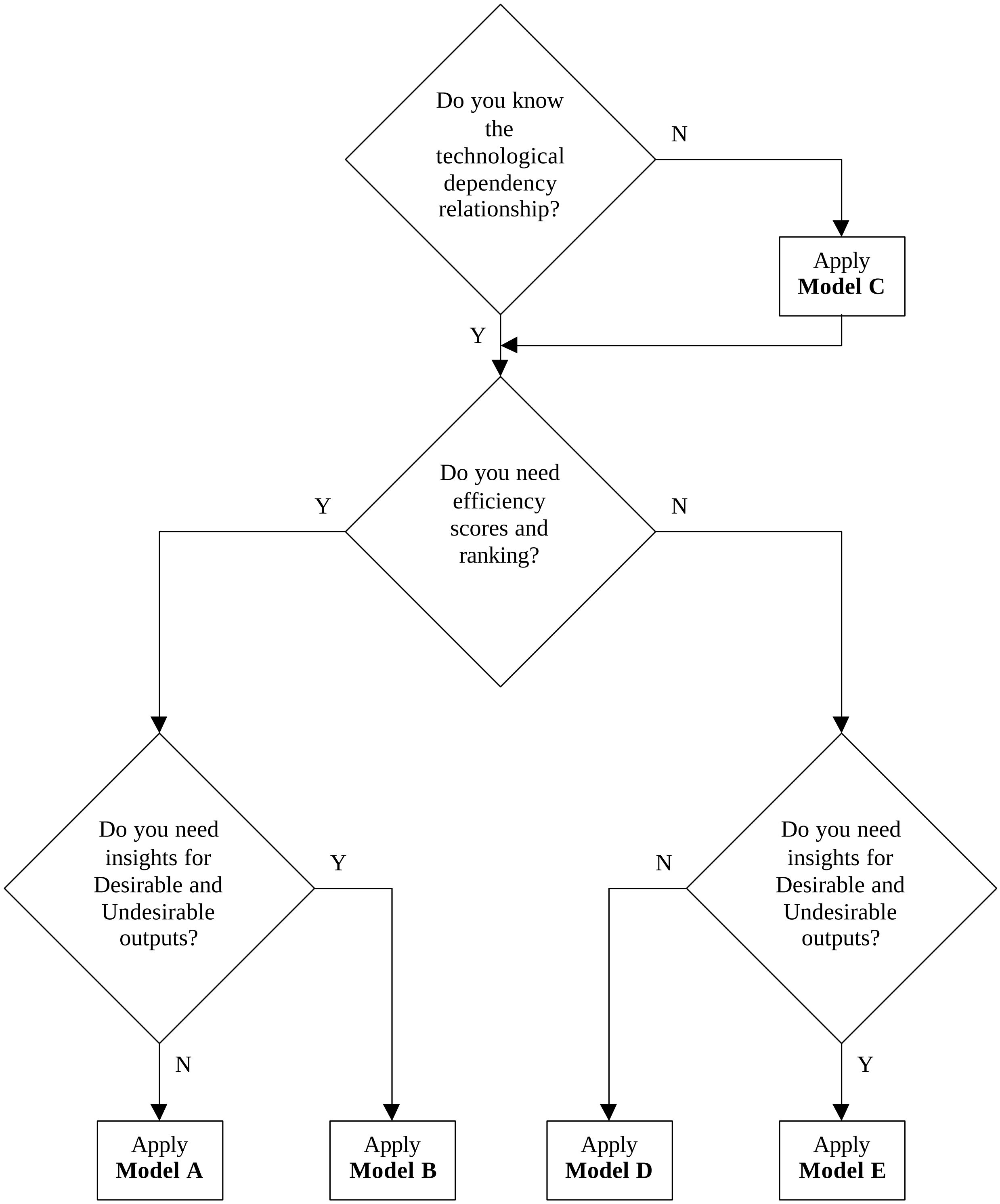
In the case of certain transformation processes, it is possible that the Decision-maker knows a priori a standard or an achievable relationship concerning the desirable and the undesirable outputs. Under these circumstances, Model A and Model B can be used by the decision-maker.

However in many instances, the technological dependence is neither known a priori nor can be calculated from the process. In such cases, Model C is used to determine the relationship between the desirable and the undesirable outputs.

Model D and Model E are applied in situations where the decision-maker needs to incorporate the relative importance to the different input, desirable and undesirable output variables and where obtaining an efficiency score is not of prime importance.

The application of the five models developed for this research has been illustrated with three different data sets. Several of these models present a different approach for efficiency evaluations in the presence of undesirable outputs. Some of these approaches- Models C, D and E are innovative. For this reason, we developed the flowchart in Figure 4.6 with the purpose of assisting the decision-maker in the selection of the most appropriate model for his/her needs.

A validation of the models through comparison with published results was not always available and or meaningful. However, whenever possible, comparisons were attempted.



Note: Model D and E need modifications to include technological dependencies

Figure 4.6. Flowchart to facilitate the selection of models

CHAPTER 5. Conclusion

In this chapter we present the conclusion of this research effort and make recommendations for future research. The first section summarizes the research of this thesis. The second section describes the major contribution of this research and the managerial implications related to the use of the different models developed. The third section outlines some recommendations for future research.

5.1. Summary

The present research study had two objectives and several sub-objectives: First, to explore new model formulations for including undesirable outputs in Data Envelopment Analysis. The goal was to find suitable formulations to address (i) a way of maximizing desirable outputs while minimizing undesirable outputs in the efficiency computations; (ii) determine performance measures for the firms in the presence of predetermined technological dependence between desirable and undesirable outputs; (iii) develop a new approach that determines the linear dependence of the desirable and the undesirable outputs and to compute efficiency scores for the firms based on their performance with respect to the two types of outputs; (v) to develop a modified model with a preferential weight structure to set targets for the firms in the presence of the desirable and the undesirable outputs and finally, to explore the Goal Programming approach to address the issue of multi-objective problems relating to inputs, desirable outputs and undesirable outputs.

A second objective was to validate and verify the suitability of the proposed models.

The undesirable outputs are an anomaly in the whole set of outputs however they are a reality given the complete production processes, with different characteristics from that of the desirable outputs. Because of this characteristic they demand a different set of assumptions related to the production possibility set and the modeling of the production process.

Furthermore, the problem of incorporating undesirable outputs into efficiency measurement requires to reward units that produce more quantity of desirable outputs and less quantity of undesirable outputs. This multi-objective nature of the problem requires the utilization of appropriate techniques to handle those differences.

The effect of managerial or policy decisions/regulations on the transformation process can be expressed in different forms according to the type of problem and system being analyzed. It can be the case that a company's policy could limit the amount of undesirable (bad) output produced even to the detriment of desirable (good output) maximization goals.

In addition, the classical DEA formulation faintly recognizes the effect of the production process on the environment or the impact that the surroundings have on the performance of the system. For the purpose of this research, the impact of the operating environment on the production process is not considered. All formulations have been developed under the assumption of a black box where the process is isolated from external influences.

This research attempts to provide ways to approach the issues described above apart from the known directions in which the problem of modeling undesirable outputs for efficiency evaluation has been explored so far in the literature. The most known approaches to handle this problem are:

1. Transformations performed on the undesirable output data so that these could be considered just as desirable outputs. The transformed undesirable output data is considered to have the same characteristics as the desirable output data. Scheel (2001) discussed various transformations that are done on the undesirable output data.
2. Other ways of including undesirable outputs relate to modifying the underlying production assumptions. Thus the two different types of outputs are incorporated differently in the formulation at the modeling stage. In this approach the data do not need to be transformed. Fare et al. (1989) defines a hyperbolic path to determine a performance measure. This type of measure attempts to increase the desirable

outputs and decrease the undesirable outputs in the same model. This approach involves non-linearity due to the hyperbolic path defined by the model. Fare et al. (1989) uses a linear approximation to take care of the non-linearity.

3. Zofio and Prieto (2001) came up with a modified model with a transformation of the same expression proposed by Fare et al. (1989) to overcome the problems with the linear approximation.
4. Fare et al. (2000) formulated an Index number approach that evaluates the performance of the firms using two separate models, one in each stage. The first stage computes an index based on the increase of the desirable outputs and the second stage based on the decrease of the undesirable outputs. Then an overall environmental index is computed using the indices from the two stages.

The drawbacks of the transformation approach are that because of the transformation done to the data set, we might run into problems associated with convexity and non-linearity. Also, in these cases the framework assumes that the transformed data has its own meaning while in many real life applications, the transformation of the data may not make sense.

There are also problems associated with Fare's linear approximation in the hyperbolic approach that are explained in section 2.7.3 of Chapter 2.

As a result of this research, the following four innovations to the DEA formulation are presented. We believe these overcome some of the problems and challenges exposed above:

1. A modification to existing formulations that considers the technological relationship existing between good and bad outputs.
2. An approach to help the decision-maker to determine this technological dependence between undesirable and desirable outputs when this is unknown along with the resulting efficiency performance based on this relationship.
3. A model with a preferential weight based structure for the desirable outputs and the undesirable outputs;
4. A GoDEA approach to consider multi-objectives for inputs, desirable outputs and undesirable outputs.

Results from all four approaches have been compared and recommendations for their potential applications have been outlined.

The significance of these innovations is clear when real production processes are observed in detail.

5.1.1. The new models: significance and insights

In the case of certain transformation processes, it is possible that the decision-maker knows a priori a standard or an achievable relationship concerning the desirable and the undesirable outputs. The decision-maker would then set this relationship as a target to be achieved. This is driven by the existence of a technological relationship/dependence between the desirable and the undesirable outputs and that he/she can explicitly determine the relationship in the form of a linear dependence. In the available literature on undesirable outputs there is a lack of discussion about the presence of any technological dependence between the desirable and the undesirable outputs.

The decision-maker would also be interested in determining the efficiency score in the presence of the interdependency and the extent to which this relation binds the production process. To take into account these issues, the previously defined technological relationship needs to be incorporated into the model in such a way that targets that are set for the non-efficient DMUs need to satisfy the technological dependence requirement. This can be done by determining hypothetical DMUs that lie to the northwest of the frontier for the purpose of comparison.

By this procedure of incorporating this new constraint, it is made sure that whenever the efficiency of a DMU is evaluated, it is compared to DMUs (real or hypothetical) on or to the northwest of the line. This step might increase the reference set (peers) of the DMUs. This is because, by adding the new constraint, we will be comparing each DMU with a hypothetical or an observed DMU that is efficient and satisfies the constraint as well. Hence, some of the non-efficient DMUs need to be travelling on a longer path to the frontier to satisfy this constraint. Due to this, the procedure might also reduce the efficiency score of some of the DMUs in the observed set.

The modified hyperbolic approach for the measurement of efficiency by Zofio and Prieto(2001) as explained in Chapter 2 looks at the expansion desirable outputs by a factor q and a contraction in the undesirable outputs by a factor q. Here we apply the dependency inequality that we developed above as a constraint to the model with weak disposability assumption only. It can similarly be incorporated in the model with the strong disposability assumption.

Thus, this linear program was formulated in its envelopment form where the scores of the individual DMUs and the weights of the peers are determined. The weights of the peers determine the extent to which each of the other DMUs are important in coming up with a hypothetical DMU on the efficient frontier for the purpose of comparison with the DMU being evaluated. It can be seen that, $q¢$ is the factor of increase of the desirable outputs and the decrease of undesirable outputs. The objective function looks at maximizing the factor $q¢$.

The index number approach developed by Fare et al. (2000) as explained earlier is a two step approach to determine the efficiency scores. Each step computes an index number. The overall environmental performance index is given by the ratio of these two indices. We modified this index by incorporating the inequality constraint into this model. The two models have the additional constraint incorporated that modifies the computed score of the DMUs accounting for their performance in the presence of the ideal relationship between desirable and undesirable outputs. The desirable output and the undesirable output quantity indices are computed with respect to a specific reference DMU and then the overall environmental performance index is given as the ratio of the desirable output quantity index to the undesirable output quantity index.

We also look at a modified approach that helps the decision-maker to determine the actual interdependency between desirable and undesirable outputs. Here we deviate slightly from the basic intuition of the transformation process and come up with a different type of efficiency score based on the relationship between the desirable and the undesirable outputs. The intuition arises with the basic assumption that a linear relationship exists between the two types of outputs. Using this assumption, a different type of efficiency score r "good output" efficiency, is defined as the ratio of observed amount of desirable output over the ideal amount of desirable output.

The ideal amount of the desirable output is compute by going back to the linearity assumption used to determine the ideal amount of desirable output using the observed amount of undesirable output. It is given by an equation that contains slope of the gradient of the relationship between the desirable and the undesirable output plus the intercept with the desirable output axis. These parameters need to be determined by the linear program. The efficiency score thus computed give us insight on how the DMU is performing with respect to the desirable and the undesirable outputs alone. This does not account for the performance of the DMU with respect to the inputs. This model looks at the performance based on only the desirable and the undesirable outputs. However, the formulation can be expanded to include inputs as well in the performance measurement.

To address the problem when considering the inputs as well we developed a model to include them in the measurement process using a preferential weight structure that will incorporate the increase of the desirable outputs and the decrease of the undesirable outputs with preferential weights attached to them. Here both the increase of the desirable outputs and the decrease of the undesirable outputs are accounted for in the same model. In addition, the model has a non-radial structure. Hence this model could be used to set targets in the context of a generalized preferential structure over potential changes to desirable-undesirable output changes.

The Goal Programming approach in DEA is a method formulated by Athanassopoulos (1995) to account for issues in the Thanassoulis and Dyson (1992) model addressed in section 2.8. The model explained in this section looks at minimizing the slacks associated with the inputs, desirable and undesirable outputs.

The structure of the objective function has the overall target of minimizing any inefficiency associated with any of the variables. The model gets to the frontier by minimizing the positive and negative deviational variables associated with the inputs, desirable and the undesirable outputs. This model is different from the rest of the models developed because, unlike other models, this model looks at improving the performance by considering the inefficiency with respect to all the three different types of variables.

As a modification to the GoDEA formulation, the dependency constraint of the desirable outputs to the undesirable outputs could be added to the formulation.

The models developed were applied to three different sets of data. The mains insights obtained are:

- ❖ The hyperbolic approach model L1 is already a constrained model, where the search path is restricted to the hyperbolic curve. When adding an additional constraint to model the interdependencies, the solution space gets further reduced. Due to this, either more DMUs fall on the frontier or their distance to the frontier gets reduced. While this feature might seem as a disadvantage, it really is an advantage since by incorporating real dependencies between good and bad output we are constraining the solutions to those DMUs that operate under this relationship.

- ❖ The original Fare's index number approach model L2, is characterized by a weak disposability assumption while by incorporating the interdependencies, Model B presents strong disposability. Furthermore, the additional technological relationship constraint envelops the data more loosely, thus decreasing the efficiency scores. When modeling for the interdependencies one can think that units that do not follow the pre-determined relationship will be penalized. The score of the DMU under analysis depends on the added constraint, which determines that the undesirable output of the hypothetical DMU should be less than or equal to the calculated value of the undesirable output. This calculation is done using the interdependency relationship. One could also use the weak disposability assumption for model B.

- ❖ Model C produces the equation's parameters —the slope and intercept —of the line that depicts the relationship between desirable and undesirable outputs. Some DMUs that in other models are deemed inefficient can become efficient when applying this model. This probably can be explained by the fact that inputs are included in Scheel's formulation, which changes the case mix and therefore, the frontier.

- ❖ By applying the preferential weights structure model D some DMUs experienced a decrease in their efficiency score while others experienced an increase as compared to Fare’s Model L2. This is due to the different mechanisms that both formulations use to arrive to the efficiency score. Fare’s model calculates the efficiency score for desirable and undesirable outputs separately, calculating an index. Therefore, the reference set for each DMU will be different for the desirable outputs as compared to the reference set for the undesirable outputs. Contrary to that, Model D, calculates the efficiency score considering both, desirable and undesirable outputs at the same time, therefore, comparing the DMU to only one hypothetical DMU. The reference set in this case needs to satisfy convexity assumptions for both, desirable and undesirable outputs at the same time, therefore, influencing the efficiency score which might be different than the one calculated in Fare’s model.

- ❖ The Goal programming formulation (Model E) does not provide efficiency scores but identifies the shortfalls and the excesses of the corresponding variables. By applying the goal programming approach of model E the slacks of the desirable output are increased for a number of DMUs while the slacks for undesirable outputs in the goal programming case are zero for more DMUs than in Model L1. The only case where the slacks are comparable is on the input side, where the slacks are very close. Similar values have been obtained for six DMUs in both models while four DMUs had zero slack in both models. Four DMUs had substantially different values.

5.2. Research Contribution and Managerial Implications

This study has contributed to the field of Data Envelopment Analysis and performance evaluation in several ways. Prior to this research, there were no models dealing with the technological dependency among desirable and undesirable outputs, even though in most production systems, such interrelationship does exist. Several approaches modeled for the presence of undesirable outputs only without considering the potential interrelationship of these with the desirable outputs. This research has solved this problem augmenting such models with a constraint that takes care of such interrelationship if it exists.

Second, traditional efficiency evaluation has considered the ratio of outputs to inputs only. In some cases, however, it is possible that the decision-maker is equally concerned with evaluating decision-making units on the basis of their output production (including desirable and undesirable outputs). For these cases, this study has proposed a different conceptualization of efficiency, the "Good output" efficiency, in which DMUs are evaluated regarding their ratio of expected good output production to the observed good output production. The expected good output production being a function of the interrelationship with bad outputs. We consider this innovative formulation as, perhaps, one of the most important contributions of the present study.

Third, in most production systems the decision-maker encounters multiple production goals: he or she wants to maximize the production of desirable or good outputs while reducing or keeping constant resource (input) consumption and reducing the production levels of undesirable outputs. To address this multiple-objective problem a goal programming approach has been proposed for the first time.

5.2.1. Managerial Implications of all models

All five formulations are different and have different managerial implications. The choice of model by decision-makers is based on the type of information at hand, the models' requirements, and the different, sometimes conflicting managerial objectives. For example, both, models A and B require that the decision-maker identifies the technological relationship between desirable and undesirable outputs. This might be the case for some industry applications concerned with not exceeding a pre-determined amount of pollutants because of environmental regulations. Hence the name “environmental performance index.” Some other industrial applications might also fall into this category, when this “ideal” relationship is known a priori, as is the case with some molding and extrusion processes where the expected ideal relationship between scrap and good usable output is part of the machine specifications.

Conversely, as articulated before, the motivation to apply model C would be to learn the technological relationship between good and bad outputs, if any. Whenever the most important aspect of production is to evaluate units producing bad and good outputs relative to bad output minimization this model is most adequate. After learning the technological dependency this model will produce an efficiency score to rank the DMUs according to this new “*output efficiency index*” based on undesirable to desirable output production. It is the case for so many producing units that the production of undesirable outputs is taken for granted, without knowing the type of interdependencies that bad and good output production might have. Once the technological relationship is defined, then DMUs are evaluated in light of the ideal ratio of desirable to undesirable outputs.

The application of Model D requires that the decision-maker be prepared to provide input into the model. For that, he or she needs to define a priori the most desirable hierarchy of preferences given to the different outputs – good and bad—so that a set of weights is attached to them in the objective function. This will give form to the preferential structure formulation. Consequently, this model is useful when management is interested in knowing the performance of the DMUs regarding a determined preference regarding the production of one or more desirable outputs against undesirable outputs and to set the ideal production targets based on that structure.

The usage of model E is constrained to those applications in which the decision-maker is faced with conflicting goals regarding input and bad output minimization versus good output maximization. This model can be used whenever the most important outcome expected is not to have an efficiency score, but rather to identify the relative best possible combinations or production targets for inputs, and bad and good outputs.

5.3. Recommendations for Future Research

The current research can be extended with respect to one or more of its components, namely, DEA, undesirable output evaluation, and goal programming in DEA (GoDEA).

With respect to DEA approaches that deal with undesirable outputs, the suitability of the path by which the frontier is approached can be an issue of interest. Moreover, the impact that the path to the frontier has on the efficiency results also warrants attention. As further research, one could experiment with other non-radial paths, the presence of regulations and standards, etc.

In this research, we developed a model to take into account "Good Output" efficiency. This model needs to be further amplified to account for input consumption as well. Developing a model in which the input, desirable output and undesirable output are combined to account for output efficiency has a potential for further research.

With respect to Goal Programming, or GoDEA, the approach needs further investigation not only to find the best way to apply it but also ways of best communicating its results to the decision maker.

On the application side, novel applications of the formulations developed in this research can be experimented with, such as in service quality arenas, where the number of complaints, or dissatisfied customers can be considered undesirable outputs. Another potential application is in social and health services with the need to model bad outcomes. Some of these variables can be in categorical format and would require further investigation to suit the formulations to these special cases.

Bibliography

Ali, A. I. and Seiford, L. M. (1990), "Translation Invariance in Data Envelopment Analysis", *Operations Research Letters*, Vol. 9, pp. 403-405

Athanassopoulos, A. D. (1995), "Goal Programming & data envelopment analysis (GoDEA) for target-based multi-level planning: Allocating central grants to the Greek local authorities", *European Journal of Operational Research*, Vol. 87, Issue 3, pp. 535-550.

Banker, R. D., Charnes, A. and Cooper, W. W. (1984), "Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis", *Management Science*, Vol. 30, No. 9, pp. 1078-1092.

Brannlund, R., Chung, Y., Fare, R. and Grosskopf, S. (1998), "Emissions Trading and Profitability: The Swedish Pulp and Paper Industry", *Environmental and Resource Economics*, Vol. 12, Issue 3, pp. 345-356.

Brannlund, R., Fare, R. and Grosskopf, S. (1995), "Environmental regulation and profitability: An Application to Swedish Pulp and Paper Mills", *Environmental and Resource Economics*, Vol. 6, No. 1, pp. 23-36.

Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (1994), "Data Envelopment Analysis: Theory, Methodology and Applications", Kluwer Academic Publishers, Boston.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research*, Vol. 2, pp. 429-444.

Charnes, A., Cooper, W. W. and Sueyoshi, T. (1988), “A Goal Programming/ Constrained Regression Review of the Bell System Breakup”, *Management Science*, Vol. 34, No. 1, pp. 1-26.

Chung, Y. H., Fare, R. and Grosskopf, S. (1997), “Productivity and Undesirable Outputs: A Directional Distance Function Approach”, *Journal of Environmental Management*, Vol. 51, Issue 3, pp. 229-240.

Coelli, T., Prasada Rao, D. S. and Battese, G. E. (1998), “An Introduction to Efficiency and Productivity Analysis”, Kluwer Academic Publishers, Boston.

Cooper, W. W., Seiford, L. M. and Tone, K. (1999), “Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software”, Kluwer Academic Publishers, Boston.

Fare, R., Grosskopf, S. and Hernandez-Sancho, F. (2000), “Environmental Performance: An Index Number Approach”.

Fare, R., Grosskopf, S. and Lovell, C. A. K. (1985), “The Measurement of Efficiency of Production, Studies in Productivity Analysis”, Kluwer-Nijhoff Publishing, Dordrecht.

Fare, R., Grosskopf, S. and Lovell, C. A. K. (1994), “Production Frontiers”, Cambridge University Press, London.

Fare, R., Grosskopf, S., Lovell, C. A. K. and Pasurka, C. (1989), “Multilateral Productivity Comparisons when some Outputs are Undesirable: A Non-Parametric Approach”, *Review of Economics and Statistics*, Vol. 71, Issue 1, pp. 90-98.

Fare, R., Grosskopf, S. and Pasurka, C. (1986), “Effects on relative efficiency in electric power generation due to environmental controls”, *Resources and Energy*, Vol. 8, pp. 167-184.

Fare, R., Grosskopf, S., Lovell, C. A. K. and Yaisawarng, S. (1993), "Derivation of Shadow Prices for Undesirable Outputs: A Distance Function Approach", *The Review of Economics and Statistics*, Vol. 75, Issue 2, pp. 374-380.

Farrell, M. J. (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, Vol. 120, Series A (General), Part III, pp. 253-278.

Golany, B. and Roll, Y. (1989), "An Application Procedure for DEA", *OMEGA: The International Journal of Management Science*, Vol. 17, No. 3, pp. 237-250.

Hoopes, B. J., Triantis, K. P. and Partangel, N. (2000), "The Relationship Between Process and Manufacturing Plant Performance: A Goal Programming Data Envelopment Analysis Approach", *International Journal of Operations and Quantitative Management*, Vol. 6, No. 4, pp. 287-310

Koopmans, T. C. (1951), "Analysis of Production as an Efficient Combination of activities. In: Koopmans, T. C. (Ed.), Activity Analysis of Production and Allocation, Cowles Commission, Wiley, New York, pp. 33-97.

Otis, P. (1999), "Dominance Based Measurement of Environmental Performance and Productive Efficiency of Manufacturing", Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Department of Industrial and Systems Engineering, Blacksburg, Virginia.

Scheel. H. (2001), "Undesirable Outputs in efficiency valuations", *European Journal of Operational Research*, Vol. 132, pp. 400-410.

Shephard, R. (1953), "Cost and Production Functions", Princeton University Press, Princeton, New Jersey.

Thanassoulis, E. and Dyson, R. G. (1992), "Estimating preferred target input-output levels using data envelopment analysis", *European Journal of Operational Research*, Vol. 56, pp. 80-97.

Zofio, J. L. and Prieto, A. M. (2001), “Environmental efficiency and regulatory standards: the case of CO_2 emissions from OECD industries”, *Resource and Energy Economics*, Vol. 23, pp. 63-83.

Appendix I. Data Tables

Zofío and Prieto (2001) Data Set for OECD Countries

S. No.	DMU	Output (Mill US$)	CO_2 (Mill. Tons)	Capital Stock (Mill US$)	Labor (Thousands)
1	CAN	265184.62	98.13	100661.54	1870.00
2	USA	2861273.00	886.25	1152700.00	17496.00
3	JAP	1664941.12	281.35	687898.78	11173.00
4	AUS	120555.04	46.18	55004.32	1015.50
5	BEL	91728.24	36.56	49979.54	779.60
6	FIN	44651.41	15.27	28702.19	434.40
7	FRA	559702.27	101.98	268013.16	4558.50
8	GER	847046.41	143.41	335110.05	7210.00
9	GRE	33378.36	10.35	24847.65	346.30
10	ITA	403049.26	83.10	313520.06	2806.80
11	NOR	32236.07	8.61	22015.83	271.80
12	SWE	89630.94	13.29	31820.29	717.80
13	UK	564430.00	94.65	290166.67	4849.20
14	SPA	241972.60	48.20	115402.49	2046.40

Scheel's (2001)Data Set for 13 Countries

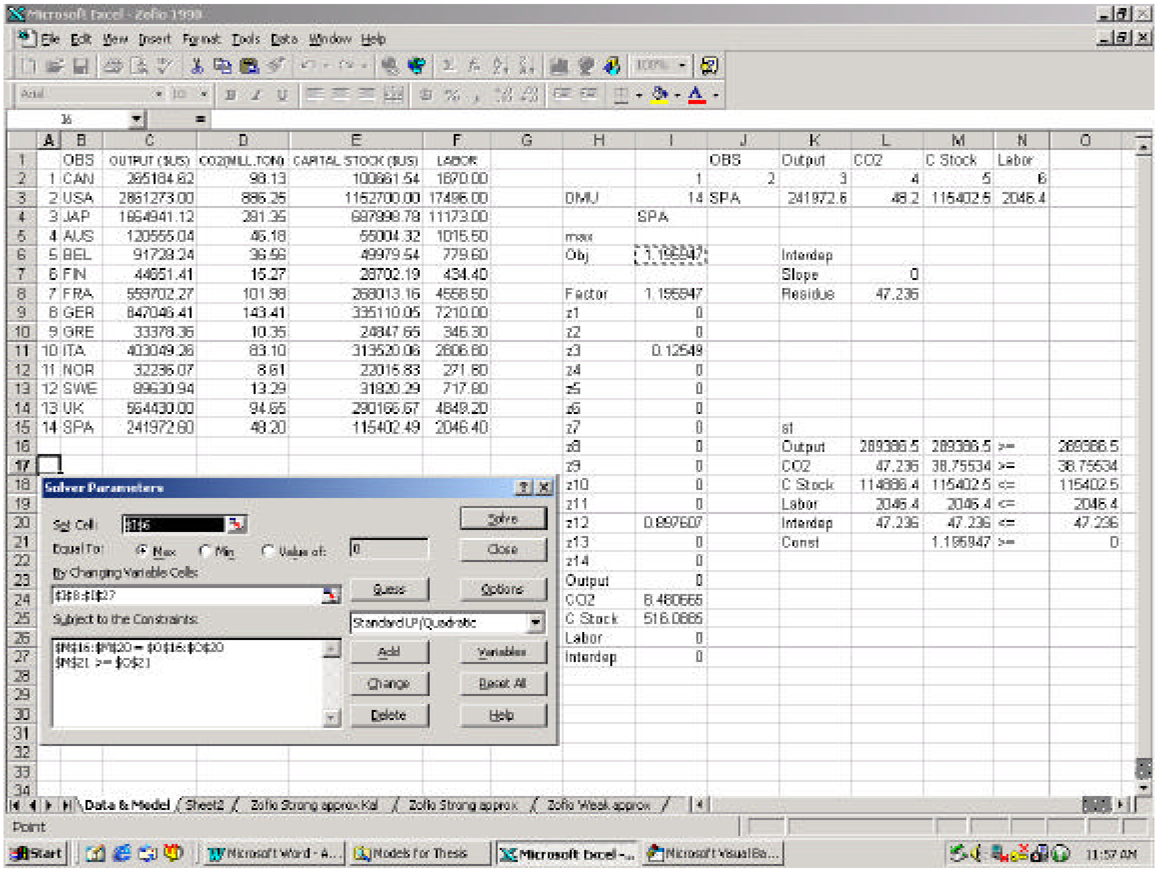
S. No.	DMU	GDP (DM)	NO (Mill. Tons)	Employees (Thousands)
1	B	385.7	76	3793
2	D	3457.4	612	35782
3	DK	248.2	61	2601
4	E	802.4	151	12027
5	F	2204.0	294	22057
6	GB	1579.3	456	25936
7	GR	163.8	33	3821
8	I	1560.9	295	19943
9	IRL	92.2	24	1262
10	NL	566.9	116	6782
11	P	144.0	24	4417
12	S	330.9	78	4134
13	SF	179.2	54	2016

The Washington Post Data Set from the Pressroom

S. No.	ID	Good Papers	Waste	Press Downtime (Min.)	Time (Min.)
1	4745	90913	690	8	99
2	4746	123960	1270	19	160
3	4747	85316	820	54	133
4	4748	120348	2600	18	142
5	4801	109740	900	44	160
6	4802	108452	1600	0	122
7	4803	79729	1300	27	114
8	4804	123062	4600	21	160
9	4857	120308	1770	18	162
10	4858	77042	1020	64	171
11	4859	98496	1380	28	152
12	4860	126844	1250	23	158
13	4913	102932	1580	24	150
14	4914	100596	1120	34	168
15	4915	127256	890	32	160
16	4916	86730	800	42	141
17	4966	85350	1130	66	151
18	4967	82581	1820	29	130
19	4968	127018	1300	18	149
20	4969	125282	1550	5	135
21	5015	92498	1440	49	166
22	5016	126587	1510	19	173
23	5017	99828	1600	42	164
24	5018	104446	950	20	130
25	5064	124389	790	61	205
26	5065	30981	2660	119	65
27	5066	118826	1080	36	167
28	5067	143875	1500	52	200
29	5115	69363	660	73	139
30	5116	85303	2260	13	99
31	5117	108832	1000	42	151
32	5118	90264	1400	46	144

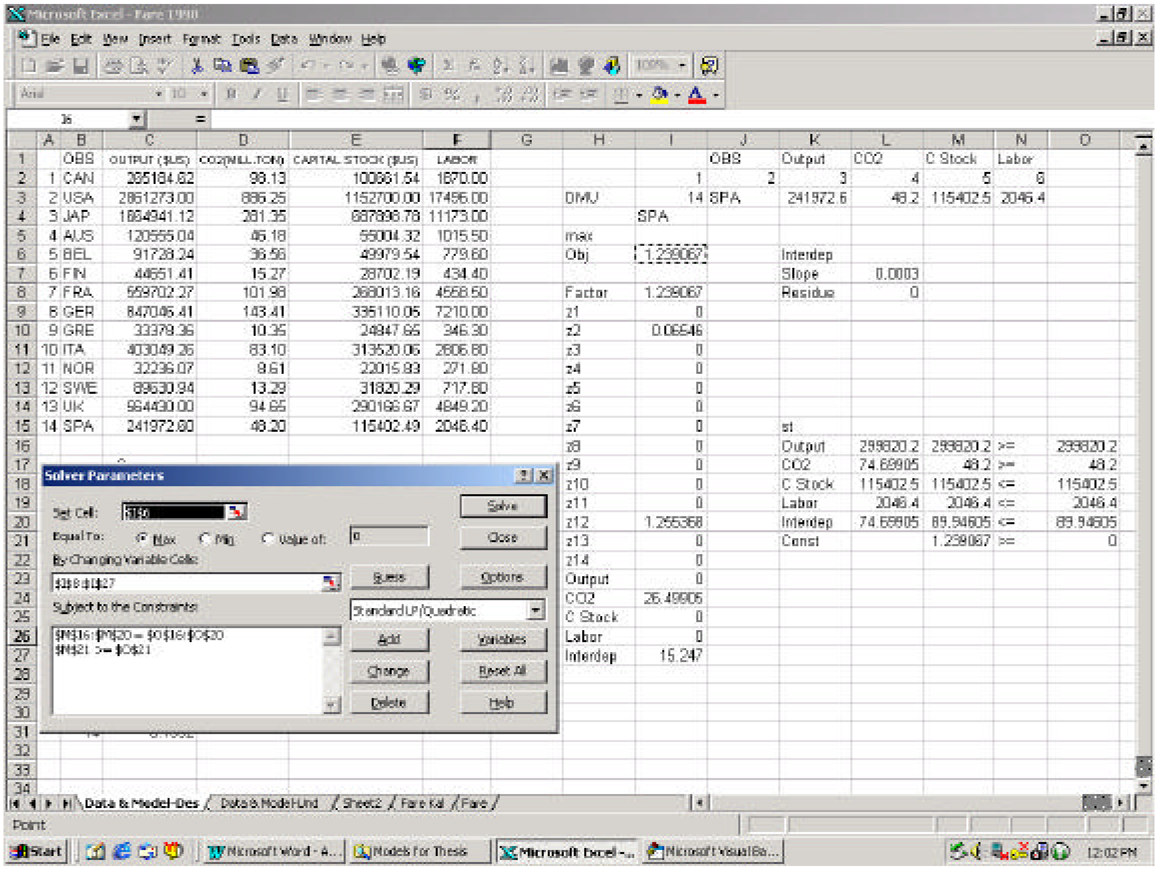
Appendix II: Excel Solver Code

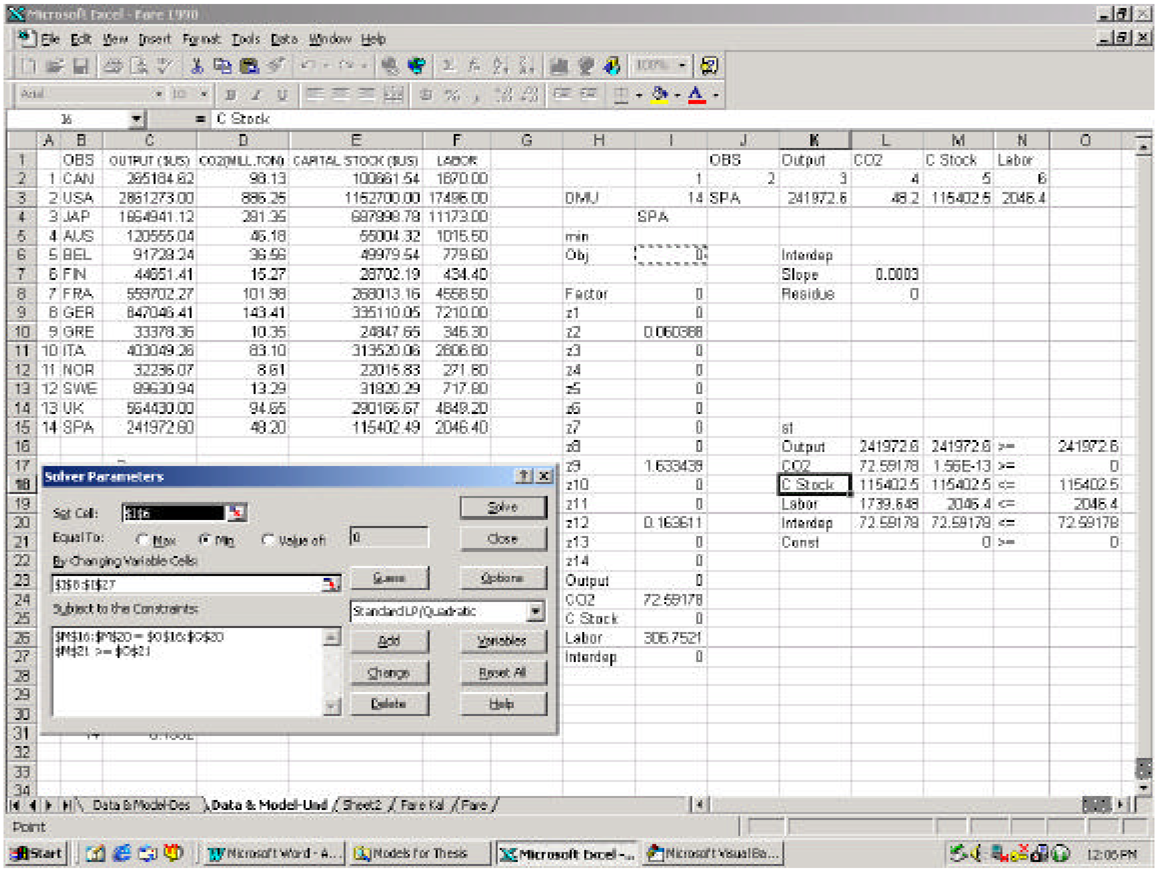
Excel Solver Code for Model A



```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 10/01/2001 by pasupatk
'
' Keyboard Shortcut: Ctrl+q
'
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 14
    Sheets("Data & Model").Select
    Range("I3").Select
    Selection.Value = n
    Range("I8").Select
    Selection.Value = 1
    Range("I9:I27").Select
    Selection.Value = 0
    SolverSolve (True)
    Range("I3:I27").Select
    Selection.Copy
    Sheets("Sheet2").Select
'    Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n).Select
    Next n
    Sheets("Data & Model").Select
    Range("K12").Select
    Sheets("Sheet2").Select
    Range("I2").Select
End Sub
```

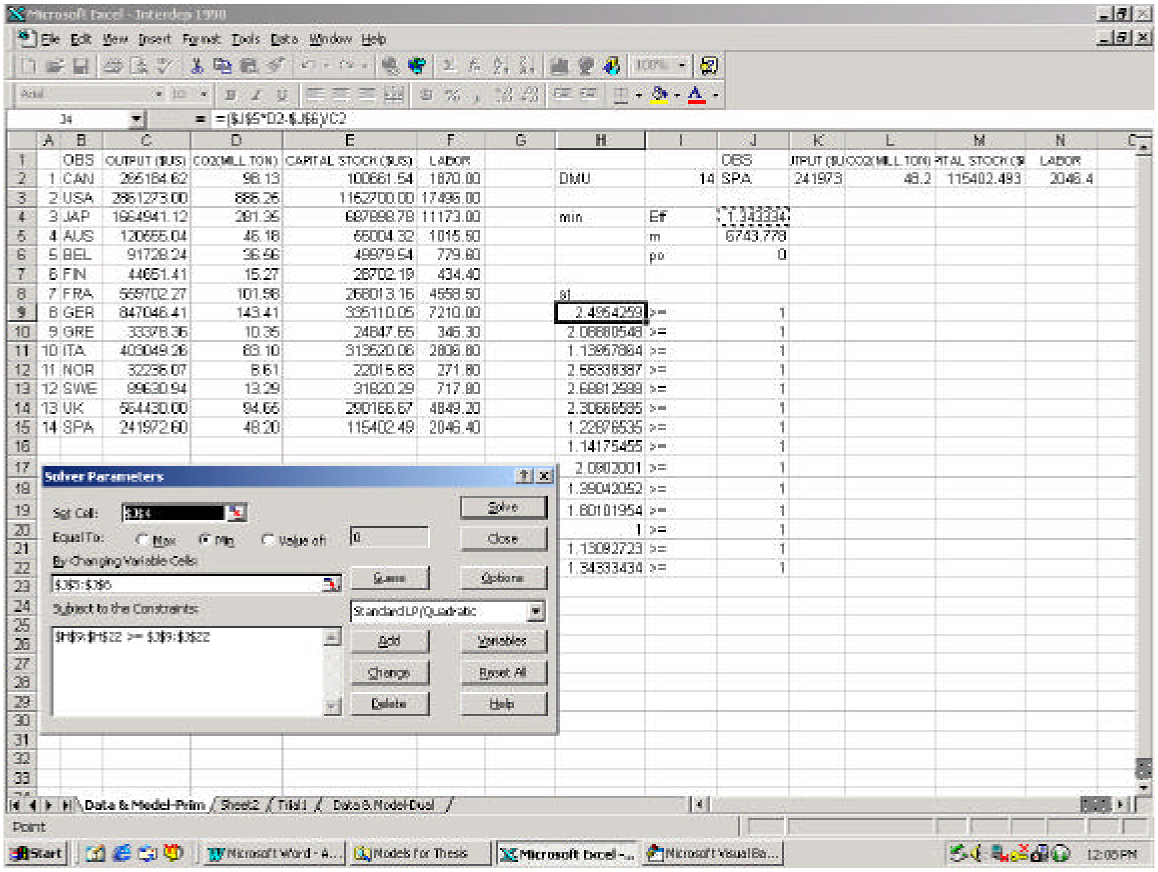
Excel Solver Code for Model B




Microsoft Excel - Fare 1990
File Edit View Insert Format Tools Data Window Help
C Stock
OBS OUTPUT ($US) CO2(MILL.TON) CAPITAL STOCK ($US) LABOR
1 CAN 285184.62 98.13 100661.54 1670.00
2 USA 2861273.00 886.25 1152700.00 17496.00
3 JAP 1664941.12 281.35 687998.78 11173.00
4 AUS 120555.04 46.18 55004.32 1015.50
5 BEL 91728.24 36.56 49979.54 779.60
6 FIN 44651.41 15.27 28702.19 434.40
7 FRA 559702.27 101.98 268013.16 4558.50
8 GER 847046.41 143.41 335110.05 7210.00
9 GRE 33378.36 10.35 24847.65 346.30
10 ITA 403049.26 83.10 313520.06 2806.80
11 NOR 32236.07 8.61 22015.83 271.80
12 SWE 89630.94 13.29 31820.29 717.80
13 UK 564430.00 94.65 290166.67 4849.20
14 SPA 241972.60 48.20 115402.49 2046.40
DMU 14 SPA
SPA
min
Obj 0
Factor 0
z1 0
z2 0.060388
z3 0
z4 0
z5 0
z6 0
z7 0
z8 0
z9 1.633439
z10 0
z11 0
z12 0.163611
z13 0
z14 0
Output 0
CO2 72.59178
C Stock 0
Labor 306.7521
Interdep 0
OBS Output CO2 C Stock Labor
1 2 3 4 5 6
SPA 241972.6 48.2 115402.5 2046.4
Interdep
Slope 0.0003
Residue 0
st
Output 241972.6 241972.6 >= 241972.6
CO2 72.59178 1.56E-13 >= 0
C Stock 115402.5 115402.5 <= 115402.5
Labor 1739.648 2046.4 <= 2046.4
Interdep 72.59178 72.59178 <= 72.59178
Const 0 >= 0
Solver Parameters
Set Cell: $I$6
Equal To: Max Min Value of: 0
By Changing Variable Cells:
$I$8:$I$27
Subject to the Constraints:
$M$16:$M$20 = $O$16:$O$20
$M$21 >= $O$21
Standard LP/Quadratic
Solve Close Guess Options Add Change Delete Variables Reset All Help
Data & Model-Des Data & Model-Und Sheet2 Fare Kal Fare
Point
Start Microsoft Word - A... Models for Thesis Microsoft Excel -... Microsoft Visual Ba... 12:05 PM

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 10/01/2001 by pasupatk
'
' Keyboard Shortcut: Ctrl+q
'
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 14
    Sheets("Data & Model-Des").Select
    Range("I3").Select
    Selection.Value = n
    Range("I8").Select
    Selection.Value = 1
    Range("I9:I27").Select
    Selection.Value = 0
    SolverSolve (True)
    Range("I3:I27").Select
    Selection.Copy
    Sheets("Sheet2").Select
'    Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n).Select
    Next n
    For n = 1 To 14
    Sheets("Data & Model-Und").Select
    Range("I3").Select
    Selection.Value = n
    Range("I8").Select
    Selection.Value = 1
    Range("I9:I27").Select
    Selection.Value = 0
    SolverSolve (True)
    Range("I3:I27").Select
    Selection.Copy
    Sheets("Sheet2").Select
'    Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n + 14).Select
    Next n
End Sub
```

Excel Solver Code for Model C



OBS		OUTPUT ($US)	CO2(MILL TON)	CAPITAL STOCK ($US)	LABOR
1	CAN	265184.62	98.13	100661.54	1870.00
2	USA	2861273.00	886.26	1162700.00	17496.00
3	JAP	1664941.12	281.36	687898.78	11173.00
4	AUS	120655.04	46.18	65004.32	1015.50
5	BEL	91728.24	36.56	49979.54	779.60
6	FIN	44651.41	15.27	26702.19	434.40
7	FRA	569702.27	101.98	268013.16	4558.50
8	GER	847046.41	143.41	335110.05	7210.00
9	GRE	33378.36	10.35	24847.65	346.30
10	ITA	403049.26	63.10	313520.06	2806.80
11	NOR	32236.07	8.61	22015.63	271.80
12	SWE	89630.94	13.29	31820.29	717.80
13	UK	564430.00	94.66	290166.67	4849.20
14	SPA	241972.60	48.20	115402.49	2046.40

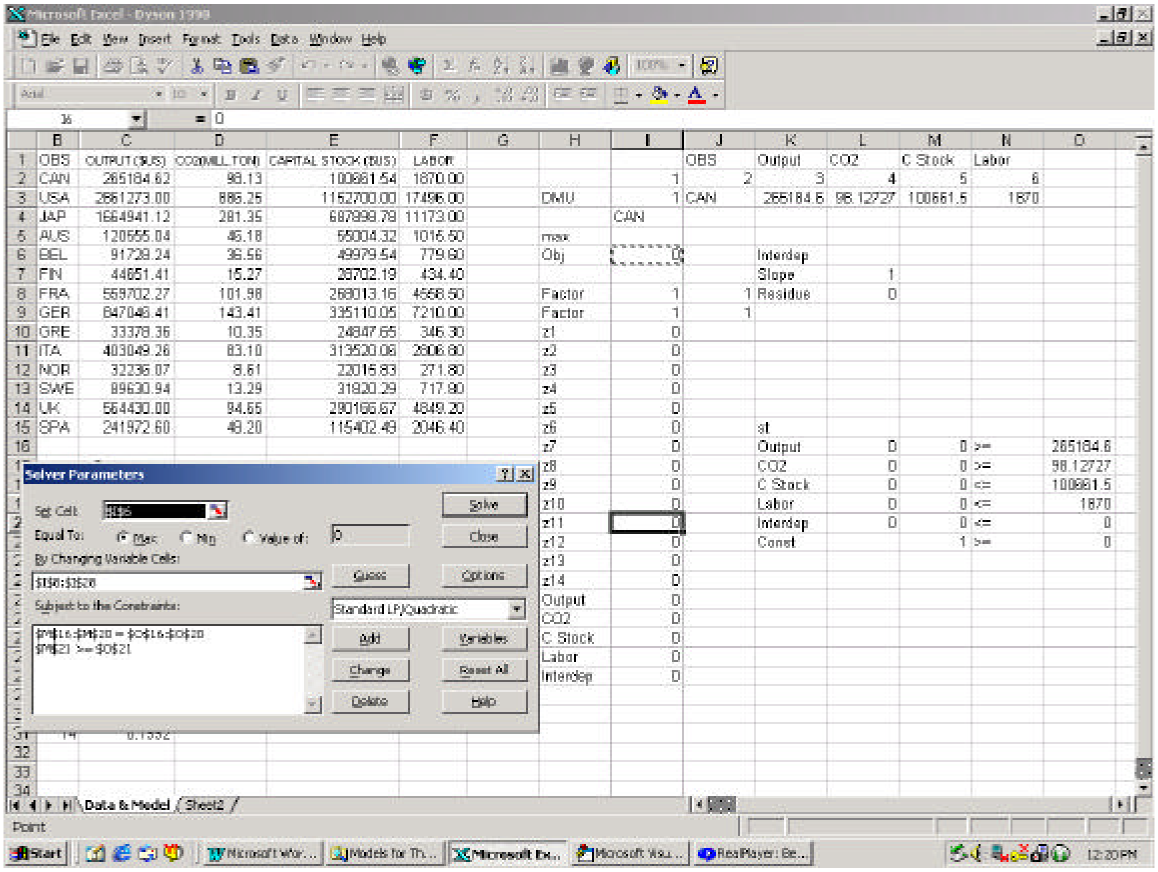
DMU	14	OBS SPA	OUTPUT ($US) 241973	CO2(MILL TON) 48.2	CAPITAL STOCK ($US) 115402.493	LABOR 2046.4

min	Eff	1.343334
	m	6743.778
	po	0

gi		
2.4964259	>=	1
2.06880548	>=	1
1.13967864	>=	1
2.58338387	>=	1
2.68812599	>=	1
2.30666585	>=	1
1.22876535	>=	1
1.14175455	>=	1
2.0902001	>=	1
1.39042052	>=	1
1.80101954	>=	1
1	>=	1
1.13092723	>=	1
1.34333434	>=	1

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 10/01/2001 by pasupatk
'
' Keyboard Shortcut: Ctrl+q
'
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 14
    Sheets("Data & Model-Prim").Select
    Range("I2").Select
    Selection.Value = n
    SolverSolve (True)
    Range("J2:J6").Select
    Selection.Copy
    Sheets("Sheet2").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n).Select
    Next n
End Sub
```

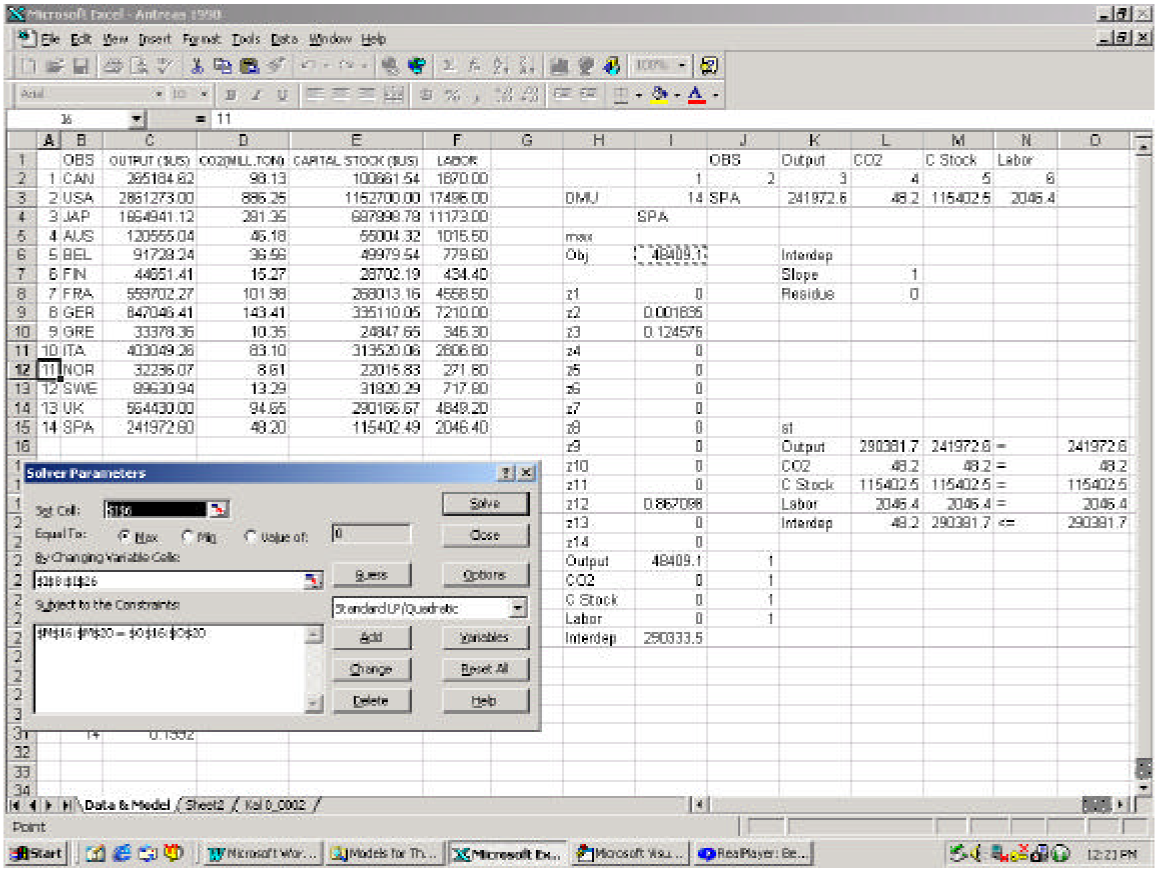
Excel Solver Code for Model D



	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	OBS	OUTPUT ($US)	CO2(MILL TON)	CAPITAL STOCK ($US)	LABOR				OBS	Output	CO2	C Stock	Labor	
2	CAN	265184.62	98.13	100661.54	1870.00			1	2	3	4	5	6	
3	USA	2861273.00	886.25	1152700.00	17496.00		DMU	1	CAN	265184.6	98.12727	100661.5	1870	
4	JAP	1664941.12	281.35	687998.79	11173.00			CAN						
5	AUS	120655.04	46.18	55004.32	1016.50		max							
6	BEL	91728.24	36.56	49979.54	779.60		Obj	0		Interdep				
7	FIN	44651.41	15.27	28702.19	434.40					Slope	1			
8	FRA	559702.27	101.98	268013.16	4558.50		Factor	1	1	Residue	0			
9	GER	847046.41	143.41	335110.05	7210.00		Factor	1	1					
10	GRE	33378.36	10.35	24847.65	346.30		z1	0						
11	ITA	403049.26	83.10	313520.06	2806.80		z2	0						
12	NOR	32236.07	8.61	22016.83	271.80		z3	0						
13	SWE	89630.94	13.29	31920.29	717.90		z4	0						
14	UK	564430.00	94.65	290166.67	4849.20		z5	0						
15	SPA	241972.60	48.20	115402.49	2046.40		z6	0		st				
16							z7	0		Output	0	0	>=	265184.6
17							z8	0		CO2	0	0	>=	98.12727
18							z9	0		C Stock	0	0	<=	100661.5
19							z10	0		Labor	0	0	<=	1870
20							z11	0		Interdep	0	0	<=	0
21							z12	0		Const		1	>=	0
22							z13	0						
23							z14	0						
24							Output	0						
25							CO2	0						
26							C Stock	0						
27							Labor	0						
28							Interdep	0						

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 10/01/2001 by pasupatk
'
' Keyboard Shortcut: Ctrl+q
'
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 14
    Sheets("Data & Model").Select
    Range("I3").Select
    Selection.Value = n
    Range("I8:I9").Select
    Selection.Value = 1
    Range("I10:I28").Select
    Selection.Value = 0
    SolverSolve (True)
    Range("I3:I28").Select
    Selection.Copy
    Sheets("Sheet2").Select
'    Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n).Select
    Next n
    Sheets("Data & Model").Select
    Range("K12").Select
    Sheets("Sheet2").Select
    Range("I2").Select
End Sub
```

Excel Solver Code for Model E



	OBS	OUTPUT ($US)	CO2(MILL.TON)	CAPITAL STOCK ($US)	LABOR
1	CAN	265184.62	98.13	100661.54	1670.00
2	USA	2861273.00	886.25	1152700.00	17496.00
3	JAP	1664941.12	281.35	697898.79	11173.00
4	AUS	120555.04	46.18	55004.32	1015.50
5	BEL	91728.24	36.56	49979.54	779.60
6	FIN	44651.41	15.27	28702.19	434.40
7	FRA	559702.27	101.98	268013.16	4558.50
8	GER	847046.41	143.41	335110.05	7210.00
9	GRE	33378.36	10.35	24847.65	346.30
10	ITA	403049.26	83.10	313520.06	2606.60
11	NOR	32236.07	8.61	22016.83	271.60
12	SWE	89630.94	13.29	31920.29	717.80
13	UK	564430.00	94.65	290166.67	4849.20
14	SPA	241972.60	48.20	115402.49	2046.40

H	I	J	K	L	M	N	O
		OBS	Output	CO2	C Stock	Labor	
	1	2	3	4	5	6	
DMU	14	SPA	241972.6	48.2	115402.5	2046.4	
	SPA						
max							
Obj	48409.1		Interdep				
			Slope	1			
z1	0		Residue	0			
z2	0.001636						
z3	0.124576						
z4	0						
z5	0						
z6	0						
z7	0						
z8	0		st				
z9	0		Output	290381.7	241972.6	=	241972.6
z10	0		CO2	48.2	48.2	=	48.2
z11	0		C Stock	115402.5	115402.5	=	115402.5
z12	0.867098		Labor	2046.4	2046.4	=	2046.4
z13	0		Interdep	48.2	290381.7	<=	290381.7
z14	0						
Output	48409.1	1					
CO2	0	1					
C Stock	0	1					
Labor	0	1					
Interdep	290333.5						

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 10/01/2001 by pasupatk
'
' Keyboard Shortcut: Ctrl+q
'
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 14
    Sheets("Data & Model").Select
    Range("I3").Select
    Selection.Value = n
    Range("I8:I26").Select
    Selection.Value = 0
    SolverSolve (True)
    Range("I3:I26").Select
    Selection.Copy
    Sheets("Sheet2").Select
'    Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Range("A1").Offset(0, n).Select
    Next n
    Sheets("Data & Model").Select
    Range("K12").Select
    Sheets("Sheet2").Select
    Range("I2").Select
End Sub
```

VITA

Kalyan Sunder Pasupathy was born in Coimbatore, India in 1978. As a school kid, he used to crave to spend time at the hill stations where his parents resided. Later he received with distinction his B.S. in Production Engineering from Bharathiar University in May, 1999. Pretty soon, he became a "Hokie" to pursue his Masters of Science in Operations Research and still cherishes a lot of small moments from the "electronic village". His summers were memorable as an intern with Menlo Logistics in the "sunny state" and with The Washington Post in Springfield, Virginia. In August 2000 he moved to Northern Virginia to conduct his research at the Systems Performance Laboratory. Kalyan joined the Ph.D. program in Management Systems Engineering in the beginning of the year 2002 and initiated his research in the Visualization of Data and Information for Decision-Making in organizations. At the moment he is working as a research assistant helping to put in place the Performance Measurement System at the American Red Cross in Washington D.C. He loves to go on long drives, spend time enjoying nature and have heated debates and endless discussions on the philosophy of life.