Chapter 2

MEASURING EFFICIENCY OF HIGHWAY MAINTENANCE PATROLS

2.1. BACKGROUND

A number of applications of DEA are found in the area of maintenance. In the particular application discussed in this chapter, we look at the performance of highway maintenance crews or patrols in the province of Ontario, Canada. The discussion herein is based on the work of Cook et al (1990), (1991), and (2001). The problem of measuring efficiency in the roadway maintenance sector is an important one that has been examined by others as well. Deller and Nelson (1991), for example, examined a similar problem but where network size is used as an output and material, labour and capital are inputs. Later, Rouse et al. (1997) revisited the road maintenance problem for the case of highways in New Zealand, by considering additional inputs and outputs. In particular, they attempt to address environmental differences among patrols by incorporating factors aimed to capture geological indicators. This was undertaken, presumably in realization of the fact that patrols are not necessarily comparable via the conventional inputs and outputs. As well, they attempt to pay attention to weight restrictions as raised earlier by Roll, Cook and Golany (1991).

At the time that the initial study of Cook et al. (1990) was conducted, the stipulated rationale for having a formal performance measure for each patrol was to permit budget setting in a resource constrained environment. As funding for maintenance has eroded over time, a need has arisen for a formal mechanism whereby patrols are treated equitably in regard to the allocation of maintenance dollars. What is most appealing about the DEA rationale in this setting, is that if an inefficient patrol can attain efficiency status, its projected inputs θX_o can aid in setting its budget, where X_o is the vector of Inputs.

Most of the routine maintenance activities on Ontario's highways fall under the responsibility of the 244 patrols scattered throughout the province. Each such patrol is responsible for some fixed number of lane-kilometers of highway, and those activities associated with that portion of the network. More than 100 different categories of operations or activities exist, and are grouped under the headings: 'surface,' 'shoulder,' 'right of way,' 'median,' and 'winter operations.'

The present system for monitoring patrol activities is the Maintenance Management System (MMS). This is a computerized record keeping system which keeps track of total work accomplished by type of operation, patrol and highway class. This system is similar to those in other Canadian provinces and states in the U.S.A.

While various statistics (such as median operations accomplished, by highway class) are maintained, there is presently no formal process for evaluating patrol activities. An area of importance to the Ministry has to do with the efficiency with which maintenance operations are carried out in various parts of the province. Since observed accomplishments influence budgetary decisions, a better understanding of efficiency will give management a yardstick for measuring what accomplishments can be expected within a given budget limit.

While there are various possible approaches to the problem of measuring efficiency in this context, the DEA framework is particularly appropriate for a number of reasons. First, the prospect of obtaining "production standards" in the usual engineering sense seems doubtful. The number of different "products" and different environmental and soil conditions mitigate against a conventional industrial engineering approach. Second, DEA is capable of handling non-economic factors, like number of accidents, maintenance dollars (an economic factor), cars/day, average age of pavement, etc., and allows for measurement of such factors on different scales. Such an approach seems particularly suited to the maintenance area, since factors such as traffic intensity, safety parameters and average age of pavements are an important part of the picture.

These and other reasons point to the appropriateness of DEA.

2.2. DEA ANALYSIS

2.2.1 The Model and its Factors

In a study of potential factors which could be utilized to best represent causes and effects relating to patrol performance, four outputs and three inputs have been chosen. Specifically, the efficiency e is given by

$$e = \frac{u_1(ASF) + u_2(ATS) + u_3(RCF) + u_4(APF)}{\upsilon_1(MEX) + \upsilon_2(CEX) + \upsilon_3(CLF)},$$

and (u_1, u_2, u_3, u_4) and $(\upsilon_1, \upsilon_2, \upsilon_3)$ denote output and input factor weights respectively.

ASF – Area Served Factor

This factor was chosen to measure the *extent* of the work load for which the patrol has responsibility. The ASF factor value is calculated from the formula

$$\text{ASF} = \sum_i \left[L_i (TLE)_i \left(A_j + C\right) + L_i \left(S_i B_j + D\right)\right]$$

where:

L_i – Length of road section *i*
TLE_i – Two-lane equivalent of road section *i*
S_i – Shoulder width of road section *i*
A_j – Coefficient for road surface type *j* (the one in road section *i*)
B_j – Coefficient for shoulder type *j* (the one in road section *i*)
C – Coefficient for winter operations
D – Coefficient for other operations (ROW, median etc.)

ATS – Average Traffic Served

This factor is intended to be a measure of the overall benefit to the users of the highway system in a patrol. The formula for computing ATS is given by

$$\text{ATS} = 10^{-3} \sum_i L_i \,(\text{AADT})_i,$$

where AADT_i is the Annual Average Daily Traffic and 10^{-3} is a scaling factor designed to bring ATS within a reasonable range for analysis.

RCF – Pavement Rating Change Factor

This factor measures the actual change in PCR, (Pavement Condition Rating) of the various road sections, relative to a 'standard' change for the same period.

APF – Accident Prevention Factor

Much of the work of maintenance staff arises due to the need to prevent accidents (surface & shoulder repairs, washouts, etc.) In this regard, accident prevention can be viewed as a cause or goal of maintenance.

A reasonable measure of accident prevention should be directly proportional to traffic level (ATS), and inversely proportional to the observed number of accidents. The chosen form is given by

$$\text{APF} = 100\frac{ATS}{C},$$

where 100 is a scaling factor and C is the number of road accidents, during the observed period, on all road sections serviced by a patrol.

MEX – Maintenance Expenditures

This is the total of all expenditures linked to the patrol. It includes both "in-house" work as well as maintenance activities performed by private contractors. Moreover, MEX includes any district-supplied services such as equipment and district supervisors' salaries.

CEX – Capital expenditures

This is the total of all capital expenditures made toward improving the existing highway infrastructure. This would include resurfacing, shoulder paving, repairs to structures, dome construction, etc. – all activities which complement maintenance efforts. Excluded are new link and new structure construction, since these do not directly complement maintenance.

CLF – Climatic Factor

What can often be an overriding consideration in the performance of a patrol, is the environmental circumstances in which that patrol must operate. The amount of snowfall, for example, will clearly influence the level of winter maintenance (snow removal and salting) needed. The extent of spring breakups will directly influence the need for summer road surface work.

Four sub-factors were taken into account in arriving at an overall climatic factor:

Snowfall
Major temperature cycles
Minor temperature cycles
Rainfall

Available data from weather stations were used to compute these sub-factors.

The overall climatic factor for a patrol is computed from:

$$\mathrm{CLF_k} = \sum_i P_{ki}(\sum_j (W_j / D_{ij})),$$

where

k – patrol index;

P_{ki} – weight of station i in calculating the climatic factor of patrol k

W_j relative importance weight of climatic factor j.

$W_1 = 50$
$W_2 = 300$
$W_3 = 20{,}000$
$W_4 = 1{,}000$

It is noted that the weights W_j were chosen while taking into account the numerical scales of each of the climatic factors (e.g. the snowfall numbers are much greater in size than the major cycle numbers). In addition, the weights were selected with attention to the resultant CLF measure being relatively of the same order of magnitude as the other efficiency factors.

2.2.2 Data and Unbounded Runs

In the present study, 4 districts are used, having a combined total of 62 patrols. As an illustration, the factor values for one of the patrols are given by:

ASF = 404
ATS = 267
RCF = 184
APF = 331
MEX = 585
CEX = 284
CLF = 715

The first level of analyses carried out uses the entire set of patrols, with 62 L.P. problems being solved. It is noted that the only constraints other than the ratio restrictions (converted to linear format) are restraints stipulating that all variables should be nonzero. This means that no patrol is permitted to assign an importance of 0 to any factor. The model is therefore, referred to as the 'unbounded' model. (The bounded model, to be discussed, will contain significant upper and lower bounds on the variables).

The results from the 62 unbounded runs are shown under column (1) of summary Table 2-1. Note the rating of 0.725 for the first patrol in District 2.

Table 2-1. Summary of Efficiencies

DMU		Efficiencies			
		Unbounded	CSW	Indiv. Wgts Entire Sample	Indiv. Wgts Within District
D	P	1	2	3	4
2	1	.725	.517	.655	.724
	3	.768	.654	.741	.878
	5	.663	.540	.614	.795
	6	.700	.606	.671	.791
	7	.650	.545	.622	.725
	9	.739	.581	.679	.751
	10	.841	.675	.756	1
	11	.948	.699	.836	1
	13	.951	.786	.912	1
	15	1	1	1	1
	16	1	.569	.761	.776
	17	1	.664	.891	.931
	18	.857	.704	.774	.839
3	1	.855	.608	.722	.988
	2	.787	.640	.744	1
	3	.756	.492	.642	.793
	4	.761	.658	.752	1
	5	1	.497	.641	.842
	6	.990	.695	.874	1
	7	1	.661	.945	1
	8	.840	.637	.811	.982
	9	.944	.624	.786	.991
	10	.613	.443	.562	.693
	11	.802	.587	.729	.923
	12	1	.528	.705	.941
	13	.921	.520	.677	.881
	14	.457	.380	.430	.608

Table 2-1 continued

DMU		Efficiencies			
		Unbounded	CSW	Indiv. Wgts Entire Sample	Indiv. Wgts Within District
D	P	1	2	3	4
8	1	.867	.528	.672	.246
	2	.885	.181	.353	.353
	3	1	.813	.940	1
	4	.998	.822	.881	.925
	5	1	.966	1	1
	6	1	.757	.849	.912
	7	.876	.745	.824	.906
	8	1	.852	.949	.984
	9	1	.727	.836	.945
	10	1	.706	.980	1
	12	1	.832	.903	.955
	13	1	.673	.869	.939
	14	1	.739	.879	.956
	15	1	.430	1	1
	16	.871	.772	.825	.883
	17	.942	.803	.872	.897
	18	1	.591	.763	.870
	19	.826	.648	.753	.885
	21	1	.724	.848	.953
	22	1	.745	1	1
	25	.817	.659	.799	.875
20	1	.583	.440	.526	.526
	2	.739	.213	.369	.370
	3	.989	.343	.541	.541
	4	1	.746	.914	.974
	5	.770	.593	.674	.674
	6	1	.836	.948	.966
	7	1	.964	1	1
	8	.781	.455	.597	.600
	9	.915	.449	.625	.635
	10	1	.710	.890	.921
	11	.819	.419	.586	.586
	12	.927	.609	.762	.768
	13	.933	.670	.795	.843
	14	1	.643	.849	.849

2.2.3 Bounded Runs

It must be emphasized again that the unbounded model yields efficiency ratings that tend to credit the patrol with a higher level of performance than may be justified. Since complete flexibility in choice of weights is permitted, the model will often assign unreasonably low or unreasonably high weights (multipliers) to some factors in the process of trying to drive the efficiency rating for the patrol in question as high as possible. Moreover, the weight assigned to a factor (e.g. CEX) by one patrol may differ drastically from the weight assigned to that factor by another patrol. Thus, in order to exercise some reasonable level of control over the manner in which importance weights are assigned, bounds need to be imposed in the model.

Given a set of absolute bounds L_i^1, U_i^1 on output multipliers and L_j^2, U_j^2 on inputs, the constraints $L_i^1 \le \mu_i \le U_i^1$ and $L_j^2 \le \upsilon_j \le U_j^2$ are added to model (1.3) of Chapter 1.

The efficiency ratings resulting from runs of this bounded version of the model are displayed in column 3 of Table 2-1. It is noted that the efficiencies obtained from the bounded runs are lower than or equal to the corresponding efficiencies arising from the unbounded analysis.

2.2.4 Deriving a Common Set of Weights

A case can be made, however, for having a Common Set of Weights (CSW). Being able to evaluate all patrols from a common reference point provides one basis for rank ordering the DMUs from best to worst. While no "best" method exists for determining such a set of weights, a simple procedure was developed for the organization in question.

Briefly, the procedure works as follows: Choose the highest priority factor, (e.g., μ_1), and while restricting all factor weights to be within their respective bounds, maximize (or minimize) the weight for the factor in question. In this particular case μ_1 is chosen as a first priority since it is both a reliable measure of output and is believed to strongly affect efficiency. The factor weight is maximized if the indicated direction is "up," and is minimized if the direction is "down."

When the optimal weight value (e.g. $\mu_1 = 800$) is determined, it is then fixed at that level in the later optimization stages. The next factor in priority is then chosen (e.g. υ_3), and minimized subject to the same constraints as applied previously, but with $\mu_1 = 800$. This process is continued until all factor weights have been set and the Common Set of Weights is established.

Efficiency ratings using the CSW are shown in column 2 of Table 2-1. Note that patrol 15 in District 2 has an efficiency rating of 1.0, when using these weights. Thus, at least in this case, the CSW is feasible.

2.2.5 District Runs

In order to extract maximum information for effective managerial control, the DEA model was run for each district separately. The resultant set of district efficiencies appears in Column 4 of Table 2-1. It is noted that these district efficiencies are higher than the corresponding values obtained when the entire set of patrols was considered. The smaller comparison groups in the district analyses give rise to this phenomenon. It is also the case that some patrols which were inefficient in the earlier analysis, obtained a rating of 1.0 in the district setting, since those efficient patrols in *other* districts against which comparison was made have been removed from the peer group.

Because significant differences may exist from one district to another (for example, climatic and highway type differences), the intra-district efficiency measures of column 4 in Table 2-1 may provide a fairer appraisal of performance. At the same time, it is desirable to detect any district-to-district differences, necessitating inter-district comparisons. Overall district performance can be viewed in a number of ways. Two useful measures that can be derived are technical efficiency and managerial efficiency.

Technical Efficiency – with this measure we compare "best" performance in a district to best performance in another district. This is taken as an indicator of the 'technical potential' of a district. Simply speaking, technical efficiency is a measure of the distance of the district frontier from the overall frontier.

One technique for obtaining this measure is to bring all points in a district to the district frontier by applying the "adjustment" method proposed in Charnes et al (1978). A somewhat simpler approach is to "correct" the district efficiencies by dividing the overall efficiency of each patrol (column 3 of Table 2-1) by the relative efficiency within the district (column 4). The resulting quotients are approximations of individual patrol efficiencies if they were brought to the district frontiers.

Taking the average of all corrected efficiencies within a district is then a measure of technical efficiency. These values are shown in column 2 of Table 2-2. it is noted, for example, that the best performance of district 20 (.986) is near the best for the entire group. District 3 on the other hand has its best performers only at 79% of the overall best performance.

Managerial Efficiency – this measure refers to the actual performance of patrols, rather than that of best performers as above. The most reasonable measure to take is the average of the actual efficiencies for the patrols in a district. Column 1 in Table 2-2 provides the average of efficiencies when the comparison group is the overall set. Column 3 is the average when the comparison group is only that set of patrols within the district. Naturally, the latter average (column 3) is larger than the former (column 1).

Table 2-2. District Efficiencies

District	# of patrols	(1) Eff. of relative to overall frontier	(2) Ave. eff. dist. front. relative to over. front.	(3) relative to district frontier
2	13	.762	.884	.862
3	14	.716	.790	.903
8	21	.847	.938	.904
20	14	.720	.986	.732

It is noted that the managerial efficiency relative to the entire group is *approximately* equal to the product of the managerial efficiency relative to the district and the technical efficiency of the district. Exact equality fails here because of the manner in which the averages are obtained.

2.2.6 Analysis of Various Characteristics

Over and above the input parameters chosen for the analysis of patrols, there are other influences (on performance) which deserve attention. These influences can be thought of as *characteristics* or circumstances which can affect the efficiency with which a patrol operates. Two particular characteristics have been chosen:

(1) % privatization
(2) traffic level

The method used to examine a given characteristic was to (1) define levels for that characteristic, (2) separate out those patrols corresponding to the various levels, and (3) do a separate analysis on each of the subgroups arising from this separation process. As an illustration, consider % privatization. Here, a particular level (for example, 10%) was chosen as the threshold separating "high" from "low" privatization. Those patrols with a percentage at or below 10% were then subject to the aforementioned analyses. This was then repeated for patrols above 10%.

The percentage of privatization is defined as the proportion of the total maintenance budget for the patrol which is utilized on privatized jobs. The proportion can be determined from the budget codes provided in the data file from which the financial information was extracted. As an example of the type of analysis which would proceed from the setting of a threshold level, the following displays the results for District 8. (See Table 2-3).

Table 2-3. Analysis by % Privatization: District 8

	Subgroup A (above 10%)		Subgroup B (below 10%)	
Patrol #	D	A	D	B
1	.7458	.7730		
2	.3530	.4560		
3			1	1
4	.9248	1		
5			1	1
6	.9120	.9255		
7			.9060	.9060
8	.9835	1		
9			.9450	.9450
10			1	1
12			.9553	.9553
13			.9384	.9387
14			.9561	.9679
15			1	1
16			.8829	.8988
17			.8975	.8992
18	.8701	.8733		
19	.8847	.8893		
21	.9529	.9789		
22	1	1		
25			.8747	1
$\sum$	7.6268	7.8360	11.3559	11.5109
Av.	.8474	.8773	.9463	.9592

Table 2-4. District 8. Sub-group A: above 10%. Sub-group B: below 10%

		Average efficiencies	
	Number of DMUs	District Analysis	Sub-group analysis
Sub-group A (high privatization)	9	.8474	.8773
Sub-group B (low privatization)	12	.9463	.9592
Total/Average`	21	.9039	

The column labeled "D" provides the overall district efficiencies which were presented earlier and have been obtained without consideration of privatization influences. When those patrols in district 8 with privatization below 10% are examined separate from the rest, different efficiency ratings result. These are displayed under column A. Note, for example, that the rating for patrol 1 rises from .7458 to .7730. Recall that the rating for a patrol when looked at in the presence of a subgroup will always be at least as high as is the corresponding "entire group" rating. The results of this type of analysis can be summarized in terms of averages, as per Table 2-4.

As a general rule, when looking at changes in average performance from the "entire district" results to the subgroup (say low privatization) results, *small* changes point to a *positive* influence of the level of the characteristic corresponding to that subgroup. For example, in the case of low privatization in district 8, the average efficiency rating of .9592 is not significantly different than the average for these patrols when analyzed relative to the entire district (.9463). This can only be explained by the fact that very few high privatization patrols were on the frontier. Thus, low privatization patrols tend to perform better than high privatization patrols since more of the former were on the frontier than was true of the latter. On the other hand, the average efficiency rating for high privatization patrols jumped from .8474 to .8773. This means that some improvement in the performance picture for high privatization patrols occurs when the efficient low privatization patrols are removed from the analysis.

As to possible inferences which one might make in the case of, say, district 8, patrols practicing a low privatization policy tend to perform on average better than is true of those with high privatization. In the case of patrol #2, for example, 0.103 points out of the total efficiency gap of .647 (=1 - .353) can be explained by privatizing out a large proportion ($\approx$ 11%) of its work.

In general, privatization impacts are different from district to district. Overall there is no conclusive evidence that privatization increases efficiency. In fact the converse seems to be true in the case of district 20.

2.3. OUTPUT DETERIORATION WITH INPUT REDUCTION

2.3.1 Theoretical versus Achievable Targets

As with many applications of DEA, implementation in the maintenance crew setting has revealed a gap between the theoretical and realistically achievable resource reduction in inefficient units. Specifically, for a given inefficient patrol, the actual input reduction $(1-\alpha)$ deemed feasible by the maintenance supervisor and geotechnical staff, who have intimate knowledge of that patrol's highway network, generally falls short of the DEA-derived $1\text{-}\theta$ for that DMU. There is a belief that below the αX_0 level, the remaining resources would not be sufficient to keep the roadway at the same standard as is currently experienced by that DMU. The general explanation for this is that the frontier units that act as peers for such inefficient units, may be operating in a more favorable environment. In the highway setting, this can mean that the frontier units may be achieving efficiency partially because highway surface conditions are superior to those of inefficient units, or that roadway sub-grade structures result in slower deterioration in the peer patrols. As well, the model of Cook et al. (1990) fails to account for certain environmental factors such as average daily temperature.

Some attempt was made in the earlier study to control for road condition, by way of a non-discretionary input, *the average pavement rating*. This rating is, however, generally not adequate to reflect the level of ongoing maintenance needed to maintain a certain standard. This rating primarily captures visible surface conditions such as extent of pavement cracking, number and severity of ruts and potholes, etc. It would not account for sub-grade depth, total pavement thickness and so on. If kept at a desirable standard, the roadway would be expected to achieve a certain life expectancy before major rehabilitation is required. If available resources are reduced below some critical point αX_o, however, a faster deterioration would result, and the expected useful lives of roads in that patrol would be reduced.

In an attempt to provide a more acceptable DEA methodology (that would be accepted by management within the transportation ministry), the earlier model of Cook et al. (1990) was upgraded to include a provision for climatic conditions. This was done in recognition of the fact that severity of snowfall clearly influences winter maintenance expenses, while the amount of rainfall impacts summer maintenance. Cook et al (1994) present an upgraded version of the earlier model that incorporates these factors, as well as a delineation between summer and winter traffic conditions. Even with this further allowance for environmental differences, however, many patrols are still unable to achieve computed performance targets, and argue that significant anomalies still exist.

Rouse et al. (1997) experienced a similar problem, and introduced a categorical variable in an attempt to address environmental differences that exist among patrols. As presented by Banker and Morey (1986), categorical variables are intended to recognize different environments in which DMUs may operate. See also Rousseau and Semple (1993). Essentially, if the setting is one where there is a single dimension (e.g. size of bank branch) according to which DMUs can be grouped, so that those in the same category are clearly comparable, then this enhanced model structure might solve the aforementioned problem of DMU anomalies. In an attempt to apply this logic in the maintenance patrol setting, however, the authors found that there was no such *single* dimension along which patrols could be ranked. For example, much of the winter and spring maintenance is a function of snowfall, temperature, temperature fluctuations, number of freeze/thaw cycles, etc. Patrols in the north do experience lower winter temperatures, thus causing pavements there to break up more rapidly than is true in *similar* patrols with more favorable temperatures. Thus, one might be tempted to categorize patrols according to temperature (or even total days of extreme cold weather). Unfortunately, it is the number of *freeze/thaw cycles* which can cause even more pavement surface damage (although geotechnical research fails to capture precisely how much more damage). It turns out to be the case that northern patrols suffer fewer such cycles than is true of patrols in more favorable temperate locations (i.e. southern patrols). One could also point to non-climate related factors, such as extent to which sub-grades under road surfaces are influenced by poor drainage conditions (e.g. swampland). A factor such as this might serve as a categorical variable as well.

The conclusion of this investigation was that categories of DMUs could be formed in several (often conflicting) ways. While it is true that more than a single categorical input can be included, meaning that a *partial* ordering of the data is possible (see e.g., Cooper, Seiford and Tone (2000)), in the present circumstances there appeared to be so many different dimensions on which DMUs could be categorized, that the model became somewhat indeterminate. This fact rendered the categorical variable approach rather inapplicable in the environment examined.

2.3.2 Enforced Input Reduction

The conventional application of DEA (for example, the VRS input-oriented model of Banker et al. (1984)), may not be appropriate in many settings for at least two reasons. First, the projection to the frontier may not be 'slackless', which will occur if a DMU is improperly enveloped. Thus, the very idea that in order to reach a projection *on* the frontier, outputs may actually have to increase, for example, renders the model rather unrealistic in a setting where the outputs are *traffic served* and *area*. Arguably, increased outputs here can mean performing a level of maintenance above that which is currently the practice, hence providing a better and more serviceable roadway for those drivers who do use it. The second, and more serious restriction of the DEA structure, is that even if one acknowledges that a radial reduction in inputs by a factor $1-\theta$ is not feasible, there is the common presumption that a reduction of a lesser amount $1-\alpha$ (where $\alpha > \theta$) will be acceptable to management. The problem here is that even if it is accepted that a given patrol cannot forfeit more resources than $(1-\alpha)X_o$, and still provide the same level of service, budget realities can deem it *necessary* to operate with *less* resources than this level dictates. Thus, budgetary reality calls for *enforced input reduction*, often beyond the αX_o critical level. Such enforced reduction in inputs is generally accompanied by *erosion of outputs*.

The important feature of the efficiency measurement exercise here, is that the measure itself is simply a means to an end. Management wishes to use such measures as a mechanism for establishing an appropriate level of maintenance funding within the province. Equally important, it wishes to gauge the impact on the highway system in the common event of under-funding. What will be the extent of the damage to the serviceability of the highway? What are the long run implications of reduced maintenance on future capital reconstruction of the highway network?

In the event where less resources are available than needed to meet standards, management's course of action would depend on the problem setting. In a bank branch situation, for example, inadequate resources, (for example branch personnel), might simply mean that there will be longer waiting times for customers, more complaints, lost accounts, and reduced sales of financial services products. In the long run, performance suffers through deteriorating sales, and overall transactions; that is, outputs decline. In the maintenance setting, inadequate resources could result in some maintenance activities being uniformly discontinued throughout the patrol area (e.g., crack sealing could be halted, roadside activities such as grass cutting might be done less often, etc.). Alternatively, management may choose to maintain the higher traffic-volume roads to standard, while sacrificing maintenance work on less important ones. Thus, on average, the serviceability, hence the output deteriorates.

The principle issue that maintenance management now faces is to obtain not only a measure of the theoretical efficiency vis-a-vis a frontier of best performing patrols, but, as well, to evaluate this against practically achievable targets. At the same time, as indicated above, management wants to assess the likely decline in roadway standards, should an inefficient patrol be required to achieve *frontier status*. Such information can aid management in setting budget targets. Specifically, reduced standards in a patrol can have long term implications for drivers (in the form of rougher roads), and for the government agency, and ultimately the taxpayer, in the form of more frequent capital expenditures prompted by shortened pavement lives. Savings in present day maintenance expenditures would, therefore, need to be traded off against accelerated resurfacing and reconstruction options.

2.3.3 Modeling Output Erosion

Let us now examine the phenomenon of output decline within the DEA context. Assume that there are n decision making units, R outputs and I

inputs, and consider the variable returns to scale (VRS) model of Banker et al. (1984) for development purposes herein. Let X_j, Y_j denote respectively the vectors of inputs and outputs for DMU $_j$. For purposes of exposition, we also assume in this section that all variables are discretionary. In the example of the following section, however, certain variables are nondiscretionary, and are treated as such.

The ratio form of the variable returns to scale model of Banker, Charnes and Cooper (1984) (BCC), is given by:

$$\begin{aligned}
&\max \frac{uY_o + \omega}{vX_o} \\
&\text{subject to:} \\
&\frac{\mathrm{uY_j} + \omega}{vX_j} \leq 1, \qquad j = 1,\ldots,n \qquad (2.1) \\
&u, v \geq 0, \quad \omega \quad \text{unrestricted}
\end{aligned}$$

The linear programming equivalents (dual and primal problems) are:

$$\begin{aligned}
&\max\ \mu Y_o + \omega \\
&\text{subject to} \qquad (2.2) \\
&vX_o = 1 \\
&\mu Y_j + \omega - vX_j \quad \leq 0, j = 1,\ldots,n \\
&\mu, v \geq 0, \qquad \omega \text{ unrestricted}
\end{aligned}$$

and

$$\begin{aligned}
&\min \theta \\
&\text{subject to:} \\
&\theta X_o - \sum_{j=1}^{n} \lambda_j X_j \geq 0 \\
&\sum_{j=1}^{n} \lambda_j Y_j \qquad \leq Y_o \qquad (2.3) \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \qquad \geq 0, \ j = 1,\ldots,n
\end{aligned}$$

As indicated above, earlier attempts to include environmental variables, and to introduce categorical inputs failed to produce targets which many patrols deemed achievable. Hence, management has tended to *adjust* DEA targets to better reflect the reality existing in certain patrols. Specifically, patrol supervisors, in collaboration with geotechnical engineers, and regional office maintenance managers, have specified what they perceive as the maximum possible input reductions $(1-\alpha_j)\%$ in respective patrols j. These values are set with the understanding that if a reduction of more than $(1-\alpha_j)\%$ in all discretionary inputs (primarily the maintenance budget) should occur in patrol j, it is claimed that outputs will begin to *erode* by some percentage γ_j. Output erosion generally means that a lower *quality* of road maintenance is being administered, as discussed in the previous section. As indicated above, the visible consequence of insufficient resources in a patrol can mean the equivalent of discontinuing maintenance on a portion of the network. To put this in context, note that the outputs we have used in the previous study are *traffic* (total users served), and *area* (roadway and roadside combined) maintained. Reduced outputs can be viewed as fewer road users receiving adequate services.

Let us assume for purposes of model development in this section, that declared expectations of output erosions are provided in good faith and represent reality. Clearly, there can be an incentive for the patrol supervisor to overstate potential output erosion, making intended budget reductions appear highly undesirable from management's perspective. There are a sufficient number of patrol-specific anomalies, such that impacts of budget reductions can only be truly estimated by the maintenance supervisor and accompanying geotechnical staff of that patrol. Hence, senior (head office) management could potentially be 'at the mercy' of patrol staff in regard to honest declarations.

One has to remember, however, that certain realities do make it rather difficult if not impossible, for patrol management to cheat in this regard. First, geotechnical staff is generally shared by several patrols, meaning that there would be little incentive to exaggerate the resource needs of one patrol at the expense of another. As well, the claims of one district supervisor must hold up to scrutiny by other supervisors who compete for the same resources. The modeling considerations discussed herein are, therefore, correct and relevant only to the extent that erosion rates reflect what will actually happen. Issues pertaining to obtaining accurate estimates of output deterioration in patrols are, thus, primarily behavioral in nature, and beyond the scope of this research.

To model the output deterioration phenomenon, refer to Figure 2-1. Note that in this simplified image of projection, with a single input and single output, inputs are reduced with no impact on outputs up to the point $\alpha_o X_o$. From that point on, outputs are assumed to radially deteriorate at a rate of γ_o

per radial percentage unit reduction in X_o, finally projecting to a level $\hat{Y}_o$ on the frontier.

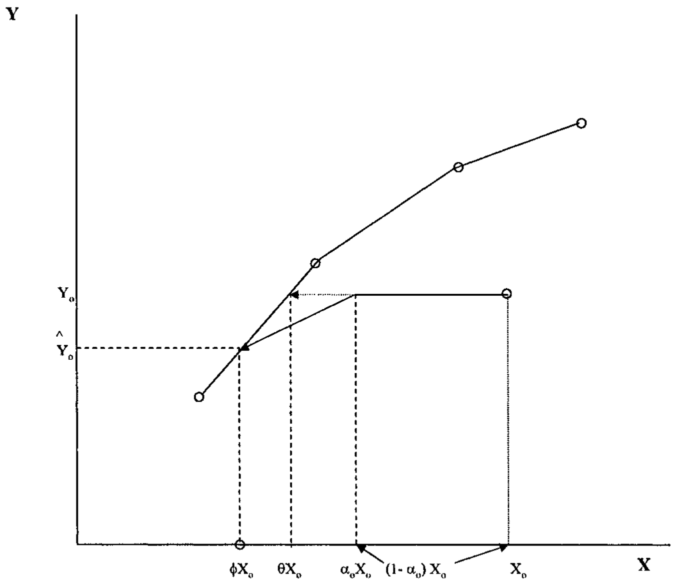

Figure 2-1. Adjusted Projection for an Inefficient DMU

In the situation studied, managers were unable to provide a precise value for γ_j. Rather, they were able to specify this parameter within bounds $\gamma_{1_j} \le \gamma_j \le \gamma_{2j}$. Estimation of the ranges $[\gamma_{1j}, \gamma_{2j}]$ posed more difficulty in some patrols than in others. Patrols with relatively uniform traffic and uniform road conditions throughout, presented less of a problem in regard to defining lower and upper bounds on γ. For those patrols where a wide range of circumstances exist among the highway sections making up its network, these ranges were, however, more difficult to capture. In this latter case one finds situations, for example, where a budget reduction can mean that a particular ditching operation to enhance drainage on a small section of the roadway may be shelved. The immediate, or even long term impact of such an activity can be difficult to quantify in terms of road deterioration, etc. Specifically, it can be the case that large budget cuts may effect few drivers, or many, depending upon the type of activity foregone. In such circumstances, management tended to specify a wider range $(\gamma_{1j}, \gamma_{2j})$ than in situations where there was more certainty. Again, we emphasize that the declared ranges are assumed to be good faith declarations, since the *zero-*

sum game environment leaves little room for any given supervisor to exaggerate his/her needs.

For model development purposes in this section we assume that γ_j is a known value. In the following section, we return to the consideration of a range $(\gamma_{1j,}\gamma_{2j})$. Let patrol o be one for which the frontier target of $(1-\theta_o)X_o$ reduction in resources is *not* achievable, but rather there is a declared maximum reduction of $(1-\alpha_o)X_o$, where $\alpha_o > \theta_o$. Formally, the primal linear programming variant (2.4) of the CCR model (2.3) becomes

$$\begin{aligned}
&\min\ \phi \\
&\text{subject to:} \\
&\phi X_o - \sum_{j=1}^{n} \lambda_j X_j && \geq 0 \\
&\sum_{j=1}^{n} \lambda_j Y_j && \geq Y_o[1-\gamma_o(\alpha_o - \phi)] \\
&\sum_{j=1}^{n} \lambda_j && = 1 \\
&\lambda_j && \geq 0,\ \ j = 1,\ldots,n
\end{aligned} \tag{2.4}$$

or

$$\begin{aligned}
&\min\ \phi \\
&\text{subject to:} \\
&\phi X_o - \sum_{j=1}^{n} \lambda_j X_j && \geq 0 \\
&-\gamma_o \phi Y_o + \sum_{j=1}^{n} \lambda_j Y_j && \geq Y_o(1-\gamma_o \alpha_o) \\
&\sum_{j=1}^{n} \lambda_j && = 1 \\
&\lambda_j && \geq 0 \quad j = 1,\ldots,n
\end{aligned} \tag{2.5}$$

While slacks are not explicitly displayed in (2.5) they do play a role in the application developed herein. More direct reference is made to slacks, and how they are computed later.

Note that the dual form of this is:

$$\begin{aligned}
&\max \mu(1-\gamma_0\alpha_o)Y_o+\omega \\
&\text{subject to:} \\
&\mathrm{vX}_0-\mu Y_o=1 \\
&\mu Y_j+\omega-vX_j \qquad \leq 0, \qquad j=1,\ldots,n \\
&\mu,\gamma \qquad \geq 0, \\
&\omega \quad \text{unrestricted}
\end{aligned} \tag{2.6}$$

and the resulting equivalent ratio model is:

$$\begin{aligned}
&\max \frac{\mu Y_o+\omega-\gamma_o\alpha_o\mu Y_o}{vX_o-\gamma_o\mu Y_o} \\
&\text{subject to:} \\
&(\mu \mathrm{Y_j}+\omega)/vX_j \qquad \leq 1, \quad j=1,\ldots,n \\
&\mu,v \qquad \geq 0 \\
&\omega \quad \text{unrestricted}
\end{aligned} \tag{2.7}$$

It is noted that since output erosion is an inherent feature in *all* DMUs, it would appear that rather than (2.7) the appropriate ratio model should be:

$$\begin{aligned}
&\max \frac{\mu Y_o+\omega-\gamma_o\alpha_o\mu Y_o}{vX_o-\gamma_o\mu Y_o} \\
&\text{subject to:} \\
&\frac{\mu Y_j+\omega-\gamma_j\alpha_j\mu Y_j}{vX_j-\gamma_j\mu Y_j} \qquad \leq 1 \\
&\mu,v \qquad \geq 0 \\
&\omega \quad \text{unrestricted}
\end{aligned} \tag{2.8}$$

It can be shown, however, that these two formulations are equivalent, as given by the following theorem.

Theorem 2.1: Problems (2.7) and (2.8) are equivalent..

Proof: It is sufficient to prove that at any point $(\hat{\mu},\hat{\omega},\hat{v})$

$$\frac{\hat{\mu}Y_j+\hat{\omega}-\gamma_j\alpha_j\hat{\mu}Y_j}{\hat{v}X_j-\gamma\hat{\mu}Y_j} \leq 1$$

if and only if

$$\frac{\hat{\mu}Y_j + \hat{\omega}}{\hat{v}X_j} \leq 1$$

Case 1: Assume $(\hat{\mu}Y_j + \hat{\omega})/\hat{v}X_j = 1$. In this case DMU j is a frontier unit, meaning that $\alpha_j = 1$. Hence,

$$\frac{\hat{\mu}Y_j + \hat{\omega} - \gamma_j \alpha_j \hat{\mu}Y_j}{\hat{v}X_j - \gamma_j \hat{\mu}Y_j} = \frac{\hat{\mu}Y_j + \hat{\omega} - \gamma_j \hat{\mu}Y_j}{\hat{v}X_j - \gamma_j \hat{\mu}Y_j} = 1$$

as well.

Alternatively, assume $(\hat{\mu}Y_j + \hat{\omega})/\hat{v}X_j < 1$.

Let θ_j denote the optimal input-oriented DEA score, for example $\theta_j = (\mu * Y_j + \omega^*)/v * X_j > (\hat{\mu}Y_j + \hat{\omega})/\hat{v}X_j$.

It follows that

$(\hat{\mu}Y_j + \hat{\omega})/\hat{v}X_j = \phi_j \leq \theta_j = (\mu * Y_j + \omega *)/v * X_j \leq \alpha_j$.

Then,

$$\frac{\hat{\mu}Y_j + \hat{\omega} - \alpha_j(\gamma_j \hat{\mu}Y_j)}{\hat{v}X_j - (\gamma_j \hat{\mu}Y_j)} \leq \frac{\hat{\mu}Y_j + \hat{\omega} - \phi_j(\gamma_j \hat{\mu}Y_j)}{\hat{v}X_j - (\gamma_j \hat{\mu}Y_j)} = \frac{\hat{\mu}Y_j + \hat{\omega} - \frac{(\hat{\mu}Y_j + \hat{\omega})(\gamma_j \hat{\mu}Y_j)}{\hat{v}X_j}}{\hat{v}X_j - (\gamma_j \hat{\mu}Y_j)}$$

$$= \frac{\hat{v}X_j(\hat{\mu}Y_j + \hat{\omega}) - (\hat{\mu}Y_j + \hat{\omega})(\gamma_j \hat{\mu}Y_j)}{(\hat{v}X_j)^2 - \hat{v}X_j(\gamma_j \hat{\mu}Y_j)}$$

$$= \frac{(\hat{\mu}Y_j + \hat{\omega})[\hat{v}X_j - \gamma_j \hat{\mu}Y_j}{\hat{v}X_j[\hat{v}X_j - \gamma_j \hat{\mu}Y_j]}$$

$$= \frac{\hat{\mu}Y_j + \hat{\omega}}{\hat{v}X_j} < 1$$

Case 2: Assume $\frac{\hat{\mu}Y_j + \hat{\omega} - \gamma_j \alpha_j \hat{\mu}Y_j}{\hat{v}x_j - \gamma_j \hat{\mu}y_j} \leq 1$.

Then $\hat{\mu}Y_j + \hat{\omega} - \gamma_j \alpha_j \hat{\mu}Y_j \leq \hat{v}X_j - \gamma_j \hat{\mu}Y_j$ or $\hat{\mu}Y_j + \hat{\omega} - \hat{v}X_j \leq (\alpha_j - 1)(\gamma_j \hat{\mu}Y_j) \leq 0$. So $(\hat{\mu}Y_j + \hat{\omega})/\hat{v}X_j \leq 1$.

Hence, the result. QED.

In the section to follow we examine the output deterioration in the context of highway maintenance crew efficiency.

2.4. THE APPLICATION

Referring again to the highway maintenance example, consider the following sample of 14 patrols.

In this example two outputs were chosen to represent the aggregate service performed by maintenance crews.

Table 2-5. Output and Input Data

	Outputs		Inputs	
Patrol#	Size	Traffic Served	Total Expenditure	Average Rating
1	696	39	751	67
2	616	26	611	70
3	456	25	538	70
4	616	31	584	75
5	560	28	665	70
6	446	16	445	75
7	517	26	554	76
8	492	18	457	72
9	558	27	582	74
10	407	18	700	69
11	463	33	630	78
12	350	88	1074	75
13	581	55	1072	74
14	413	24	696	80

Outputs

Size - a measure that is an aggregate or composite of the number of kilometres of paved surface, amount of paved versus gravel shoulders, etc.

Traffic Served - this measure accounts for the average daily traffic and the length of the roadway served.

Two inputs were used in the analysis, namely:

Inputs

Total Expenditure - the annual maintenance budget for the patrol.

Average Pavement Rating - this is a standard indicator per road section (on a 0-100 scale).

Arguably, one might consider treating *average pavement rating* as an ordinal rather than cardinal variable. In this instance, the model of Cook et al. (1993) might aid in deriving projections. It should be pointed out, however, that the rating is established through formal geotechnical data gathering and as such should be treated as quantitative rather than qualitative. With the inherent lack of precision in this measure, a somewhat more formal treatment could involve the imprecise DEA arguments of Cooper, Park and Yu (1999) and Zhu (2003;2004). We have not undertaken this herein. For a more full description of these factors, see Cook et al. (1990).

It is noted that on the input side, the available budget (total expenditure) is clearly a *discretionary* variable, while the road condition, an indicator of the environment in which the patrol operates, is clearly non-discretionary. Arguably, surface maintenance expenditures such as the filling of potholes and sealing of cracks do have a minor impact on the pavement rating (causing it to increase slightly). However, it is not really at the discretion of management to change the pavement condition in any direct way.

As discussed above, the initial analysis of patrol efficiency was conducted here for two primary reasons. First, there was a desire to determine the benchmark crews against which inefficient ones could be evaluated. This provided management with the best and, even more importantly, the worst performers, hence isolating areas where waste existed, and improvements were possible. A second, and related reason for the analysis, was to have a set of measures that could potentially aid in budget setting. Specifically, under various overall provincial highway maintenance budget scenarios, how should allocations to individual patrols be made?

The input-oriented DEA model of Banker et al. (1984) was applied, but restricting the input variable Average Pavement Rating to be nondiscretionary. Specifically, the mixed discretionary/nondiscretionary version of model (2.3) was applied, namely

$$\begin{aligned}
&\min \theta \\
&\text{subject to:} \\
&\theta x_{io} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^1 = 0, && i \varepsilon DI \\
&x_{io} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^1 = 0, && i \varepsilon NDI \qquad (2.9) \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^2 = y_{ro}, && r \varepsilon DO \\
&\sum_{j=1}^{n} \lambda_j = 1
\end{aligned}$$

Here, the set of discretionary inputs DI is the budget, and the nondiscretionary inputs, NDI consists of the single variable pavement rating. Outputs are assumed to be discretionary (DO) to the extent that under budget reductions, patrol crews can choose to service the road network in a manner that is below standard. It is noted that we explicitly represent input and output slacks here as s_i^1, s_r^2 respectively. In solving (2.9) we use a 2-stage process wherein the sum of slacks is minimized in stage 2. This unconventional way of handling slacks has some practical merit here in that for example on the input side we are identifying a minimal reduction in resources needed to reach the frontier proper from a frontier extension point.

Table 2-6 presents the projections and efficiency score θ for each of the 14 DMUs. When positive slacks exist they are displayed in brackets. In this example, exactly 7 of the patrols are efficient, both in the radial sense $(\theta = 1)$, and in the CCR-efficient sense, in that all slacks are zero (see Cooper, Seiford and Tone (2000)). The remaining inefficient units are a mix of properly enveloped (DMU#5), and improperly enveloped units (DMUs #7,9,10,11,13,14).

Table 2-6. Efficiency Scores & Projections

DMU	Size	Traffic	Expenditure	Rating	Score
1	696	39	751	67	1
2	616	26	611	70	1
3	456	25	535	70	1
4	616	31	584	75	1
5	560	28	588	75	.883
6	446	16	445	75	1
7	517	26	531	72 (4)*	.958
8	492	18	457	72	1
9	558	27	543	73.75 (.25)	.934
10	536 (129)	29.67 (11.67)	609	69	.870
11	463	33	589	72.67 (5.33)	.935
12	350	88	1074	75	1
13	581	55	855	69.75 (4.25)	.797
14	479.8 (66.8)	24	510	72.26 (7.74)	.733

**Numbers in brackets represent positive slacks. Note, for example, that the road rating for patrol #7 was 76 meaning that a projected value of 72 leaves a slack of 4.*

In attempting to apply the recommended expenditure reductions arising from the efficiency analysis, some (inefficient) patrols found that the projected values could not be achieved. In consultation with head office maintenance management, patrol supervisors provided a minimum budget level that they believed was necessary to maintain the network at a standard, as set by the department. In the case of patrol #5, for example, it was estimated that at most an 8% budget reduction was possible. Beyond this, it was felt that a reduction in maintenance effort would need to occur, and a lower quality of service would be the consequence.

As discussed earlier, an attempt was made to estimate the range $(\gamma_{1j}, \gamma_{2j})$ for the parameter γ_j, for each patrol j. Figure 2-2 illustrates how the erosion projection of Figure 1 might now appear.

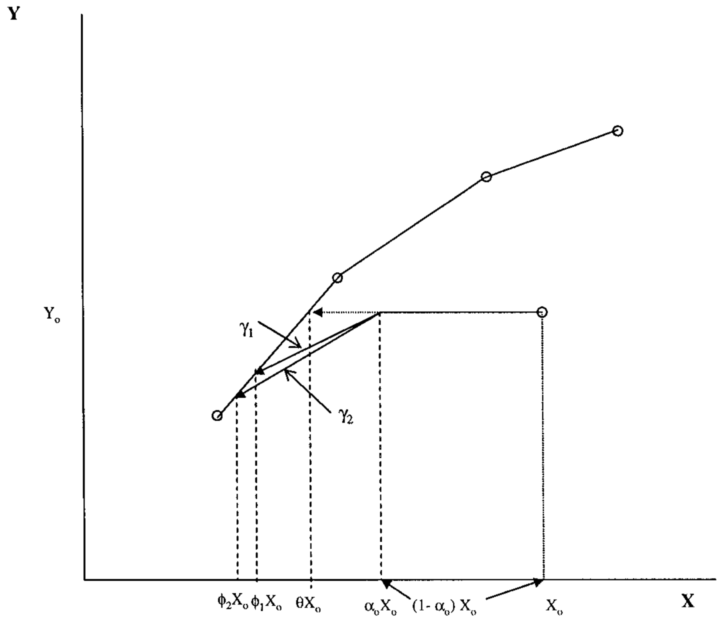

Figure 2-2. Range of Adjustable Projections

Recall that $1-\gamma_j$ is the expected percentage reduction in outputs (service) per radial percentage unit reduction in those discretionary inputs of X_j (i.e., the maintenance budget for j). For discussion purposes here, this range was taken to be $\gamma_j \in [.2,.8]$ for each j. The results for model (2.5) for each of $\gamma_{j1} = .2$ and $\gamma_{j2} = .8$, are displayed in Table 2-7. It is noted that only results for inefficient units are shown since all projections for efficient units are, by definition, the same as their current positions.

For slackless projections such as is the case for DMU #5, projected outputs help to reveal the extent of erosion of the system. Here, under the current status, size and traffic managed are represented by the values (560,28). The computed efficiency score for this patrol is .883, meaning that a reduction in expenditure of 11.7% would be needed in order to reach the frontier of best performance. The projection corresponding to this rating is shown in the row labeled 'Unadjusted.'

In this case, the claimed maximum reduction possible, without eroding outputs, is 8% (α_0 = 92% as compared to θ = 88.3%). Below the 92%

level, if outputs decline at a rate of $\gamma_1 = .2$ (20% of the input reduction beyond that point), then the resulting projected size and traffic that can be serviced are (555.4, 27.8). This represents a .7% decrease in service. Note that the new efficiency score is given by ϕ = .879. The corresponding projection for γ_2 = .8 is (527.7, 26.5), or a 5.7% decrease in outputs, with ϕ=.852. Again, see Figure 2-7.

Table 2-7. Efficiency Scores, Unadjusted and Adjusted Projections

DMU	Status	Size	Traffic	Exp.	Rating	Effic.	α
5	current	560	28	665	70	—	—
	unadj.	560	28	587.4	70	.883	—
	γ_1=.2	555.4	27.8	584.2	70	.879	.92
	γ_2=.8	527.7	26.5	566.8	70	.852	.92
	θ bdd.	543.6	27.2	587.4	70	.883	
7	current	517	26	554	76	—	—
	unadj.	517	26	531	73(3)	.958	—
	γ_1=.2	515.7	25.9	530.2	73(3)	.957	.97
	γ_2=.8	508.9	25.6	526.6	72.9(3.1)	.951	.97
	θ bdd.	512.2	25.8	530.9	73.9(2.1)	.958	
9	current	558	27	582	74	—	—
	unadj.	558	27	543.3	73.8(.2)	.934	—
	γ_1=.2	556	26.9	542.2	73.7(.3)	.932	.95
	γ_2=.8	545.8	26.4	537.0	73.5(5)	.923	.95
	θ bdd.	550.6	26.6	543.3	74.0	.934	
10	current	407	18	700	69	—	—
	unadj.	536(129)	29.7(11.7)	609	69	.87	—
	γ_1=.2	536(136)	29.7(12)609	609	.87	.95	
	γ_2=.8	536(155)	29.7(12.8)	609	69	.87	.95
	θ bdd.	536(155)	29.7(12.8)	609	69	.87	
11	current	463	33	630	78	—	—
	unadj.	463	33	589.3	72.7(5.3)	.935	—
	γ_1=.2	463	33	589.3	72.7(5.3)	.935	.935
	γ_2=.8	463	33	589.3	72.7(5.3)	.935	.935
	θ bdd.	463	33	589.3	72.7(5.3)	.935	
13	current	581	55	1072	74	—	—
	unadj.	581	55	854.8	69.7(4.3)	.797	—
	γ_1=.2	573.1	54.3	839	70.5(3.5)	.783	.85
	γ_3=.8	500	47.3	724.2	73.2(.8)	.676	.85
	θ bdd.	556.6	52.7	854.8	73.2(.8)	.797	
14	current	413	24	696	80	—	—
	unadj.	479.8(66.8)	24	509.9	72.3(7.7)	.733	—
	γ_1=.2	481(78.3)	23	504.6	72.2(7.	.725	.85
	γ_3=.8	486(124.1)	21	483.6	72.1(7.9)	.695	.85
	θ bdd.	436(62)	21.8	509.9	74.6(5.4)	.733	

Thus, under the worst case scenario, patrol 5 could experience a 5.7% decrease in service delivered to the road user and to the tax-paying public. Recall that while decreased service can take several forms, it is useful to view this scenario as portraying a lower quality product, a faster deterioration of the network, and a higher capital expenditure in the long run.

For projections with slack on the output side, a slightly different interpretation takes place. Consider the two situations portrayed by patrols #10 and #14. For #10, the projected outputs are the same under all three scenarios (unadjusted, γ_1 and γ_2). For example, the frontier projected size is 536 in all situations, and the efficiency score remains at 87%. The actual projected point (on the frontier extension)is, however, given by

$$
\begin{aligned}
&\text{Frontier projection-slack} \\
&= 536\text{-}129 \quad = 407 \text{ in unadjusted case} \\
&= 536 - 136 \quad = 400 \text{ in } \gamma_1 \text{ case} \\
&= 536 - 155 \quad = 381 \text{ in } \gamma_2 \text{ case.}
\end{aligned}
$$

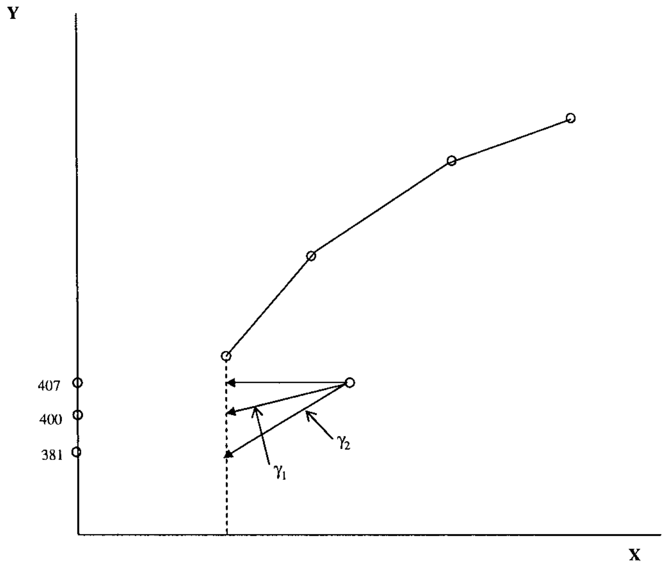

Figure 2-3. Projection with Output Slack

Figure 2-3 provides a representation of this phenomenon. (Note that a similar result occurs for the traffic factor). From a lost service perspective, it is these frontier extension values that are of interest to management.

For #14, the situation is very similar except that there is slack in only one of the outputs (size), and the efficiency score continues to decrease as we move from the unadjusted projection where θ =.733, to γ_1 (ϕ = .725) and to a_2 (ϕ = .695).

2.4.1 Base-Line Budget Considerations

The rationale for deriving input-oriented efficiency measures in the present setting, appears to be twofold. First, the measures point to those patrols that are inefficient, and those that are efficient; this sets out benchmarks that management can utilize to help poorly performing patrols to improve their status. Second, efficiency measures can aid in setting budgets. Budget planning here would appear to be an exercise in *scenario* analysis, and the results obtained from Tables 2-6 and 2-7 put bounds on the *minimal* fiscal requirements for the maintenance function. One scenario is that provided by the achievable projections described by the α_j measures. Specifically, the $(1-\alpha_j)$ % reduction in discretionary inputs (maintenance expenditure, in this case), can be achieved without any erosion to output measures. Under this scenario, for the sample of 14 patrols considered, the current budget of \$9359 could be reduced to \$8874. Thus, a budget reduction of \$485 (thousand) would appear to be immediately achievable.

The minimal budget projections under the γ_1 and γ_2 output erosion scenarios are given by \$8656, and \$8494 respectively. These lower anticipated budgets, depending on the outcome erosion rates that may result, provide management with a guide as to the possible savings obtainable if all DMUs were required to move to a frontier efficiency status.

Possibly, a more realistic and fair system of minimal budget setting would be one wherein patrols are required to reduce expenditures only by the original $1-\theta$ measure. Specifically, if no output erosion occurred, an inefficient patrol 0 would need to operate only at an expenditure level of θx_{10}, to be deemed efficient, rather than at the often lower level of ϕx_{10}. Here, x_{io} denotes the expenditure level (i=1) for DMU o. For example, in the case of patrol 13, the budget allocation would be .797 $\times$ 1072 = \$854.8 (thousand), rather than the lower figures \$839 and \$724.2 corresponding to γ_1 and γ_2, respectively. To compute the output erosion corresponding to this more favorable $0ax_{10}$ position, we resolve a modified version of (2.5) wherein ϕ is restricted to not be less than θ. Figure 2-4 illustrates this idea.

The resulting projections are shown in Table 2-7, corresponding to the status entitled θ-bdd. In computing these projections the most pessimistic view of output deterioration has been assumed (γ_2=.8 was used). Except in cases #10 and #11, the projections for inefficient units are not on the frontier, but such units would be operating at budget levels that would normally be seen as more appropriate than those resulting from the γ_1, γ_2 scenarios. The overall minimal budget for the 14 patrols in this case is \$8682. Let us regard this as a *base-line* or starting budget position.

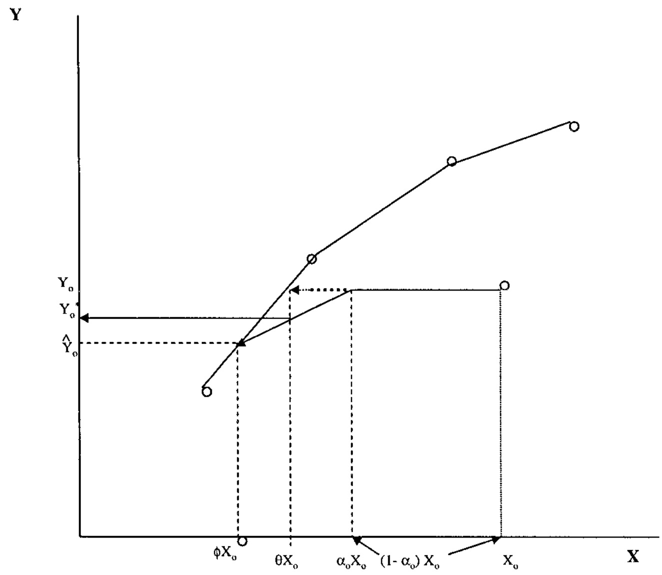

Figure 2-4. θ-Projection for an Inefficient DMU

2.4.2 Budget Allocation Beyond the Base Line

The various projections discussed above provide management with a broad scope for making budget decisions. Let us assume that the supplied γ_j (or expected γ_j) represent reality and are not exaggerated claims by the management of DMU $_j$. If the organization adopts the θx_{io} position as a form of base-budget status, then the aggregate base budget operating level is

$$B_b = \sum_{j=1}^{n} \theta_j x_{ij} \tag{2.10}$$

At this base budget level, patrol j would be providing a level of service of

$$y_j^* = y_j[1 - \gamma_j(\alpha_j - \theta_j)], \tag{2.11}$$

if j is experiencing output erosion (i.e., $\theta_j < \alpha_j$). Otherwise, $y_j^* = y_j$.

One advantage of adopting a base-budget approach as the starting point for allocating maintenance funding to patrols, is that it becomes somewhat transparent as to what budget impacts will be for funding *above* the base level. For example, if there is a $1 (thousand) increase in patrol js budget above the θx_{ij} level, one can estimate the increase in y_j that can be expected to occur. Specifically, the improved values of the components of y_j, currently at y_j^*, are given by:

$$\begin{aligned} y_j' &= y_j[1-\gamma_j(\alpha_j-\theta_j\frac{1}{x_{1j}})] \\ &= y_j^* + \frac{\gamma_j y_j}{x_{1j}} \end{aligned} \tag{2.12}$$

Note that y_j/x_{1j} is the vector of existing output rates (outputs per monetary unit of budget), and γ_j is the expected rate of increase in outputs per monetary increment to the base budget.

In the single output case (y_j is a scalar), one could allocate additional resources to patrols according to the per unit *gain factor* $\gamma_j y_j/x_{1j}$ (ranked in descending order). Specifically, if DMU j_1 has the highest gain factor, then one would presumably increase patrol j_1 s budget by δ_{j1} so that

$$\delta_{j1}\gamma_{j1}y_{j1}/x_{1j1} = y_{j1} - y_{j1}^*$$

or

$$\delta_{j1} = [(y_{j1} - y_{j1}^*)x_{1j1}]/(\gamma_{j1}y_{j1}) \tag{2.13}$$

If resources still remain, allocate funds accordingly to the patrol j_2, whose gain factor is ranked in second place, and so on.

In the multiple output case, optimization is problematic in that the patrol most desirable for a funding increment in regard to the system size dimension, may not rank highest on the traffic dimension. Thus, the problem is multi-criteria in nature, with a ranking of the patrols being available for each output type. Since the units that define the outputs are not comparable, one reasonable mechanism for ranking the patrols (for consideration for budget increments) would be to replace the vector y_j by the weighted aggregate output $\mu_j y_j$, where μ_j is the optimal multiplier vector (shadow prices from (2.9) for problem j).

Pure optimization here may be somewhat elusive in that γ_j, as discussed earlier, is known only within a range $(\gamma_{1j}, \gamma_{2j})$. Management would need to choose an appropriate value γ_j in this range if a comparison of patrols is to be made.

2.5. DISCUSSION

This chapter has examined the application of DEA in the area of highway maintenance. It has illustrated as well, the difficulty of matching theoretical and achievable targets.

The suggested modifications to the conventional DEA model help to capture the *consequences* on the output side that can occur when inputs are reduced according to the computed performance measures. The failure to realize projected reductions in resources without such consequences in many real world settings can, in most instances, be attributed to factors not included in the modeling exercise. These factors commonly pertain to the *environment* that one DMU may face versus that of its peers. This environment may be physical (differences in road sub-surface structures in maintenance patrols, for example,) or demographic (e.g., customer mix characteristics in financial services settings). Another explanation relates to the *random* nature of outputs or input requirements. In the maintenance crew setting, annual maintenance needs on highways (i.e., budget requirements) are greatly a function of weather, severity of winters, and so on. Geographical location plays an important part. It can be that frontier DMUs are those located in geographically favorable settings, where winter maintenance needs are minimal and roadway deterioration is less prevalent than in other areas. Thus, maintenance needs are random and frontier DMUs can be outliers at the lower tail of the maintenance cost distribution.

Earlier attempts to introduce categorical variables to permit comparison of a DMU to only those others that are proper peers, did not seem to resolve or explain the gap between theoretical and achievable targets. This necessitated the application of model (2.5). This model will hopefully provide a useful enhancement to the existing DEA methodology. It provides a bridge between theoretical performance targets and the practical situations facing DMU management.

REFERENCES

1. Banker, R.D., A. Charnes and W.W. Cooper (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30, 1078-1092.
2. Banker, R.D. and R.C. Morey (1986), The use of categorical variables in data envelopment analysis, *Management Science* 32, 1613-1627.
3. Charnes, A., W.W. Cooper and E.L. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research* 2, 429-444.
4. Cook, W.D., Y. Roll, and A. Kazakov (1990), A DEA model for measuring the relative efficiency of highway maintenance patrols, *INFOR* 28, 113-124.
5. Cook, W.D., Y. Roll, L. Seiford and A. Kazakov (1991), A data envelopment approach to measuring efficiency: case analysis of highway maintenance patrols, *Journal of Socio-Economics* 20, 83-103.

6. Cook, W.D. and J. Zhu (2001), Output deterioration with input reduction in data envelopment analysis, *IIE Transactions*, 35, 309-320.
7. Cook, W.D., A. Kazakov, and Y. Roll (1994), On the measurement and monitoring of relative efficiency of highway maintenance patrols, Data Envelopment Analysis: Theory, Methodology and Applications, Charnes, A., Cooper, W.W., Lewin, A. and Seiford, L. (eds.), Kluwer, New York. NY.
8. Cook, W.D., Kress, M., and Seiford, L.M. (1993), One the Use of ordinal Data in Data Envelopment Analysis, *Journal of Operations Research Society* 44, 133–140.
9. Cooper, W.W., Park, K.S., and Yu, G., (1999), IDEA and AR-IDEA: Models for Dealing with Imprecise Data in DEA, *Management Science* 45, 597–607.
10. Cooper, W.W., Seiford, L., and Tone, K., (2000), Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software, Kluwer Academic Publishers, Boston.
11. Deller, S.C. and C.H. Nelson (1991), Measuring the economic efficiency of producing rural road services, *American Journal of Agricultural Economics* 73, 194-201.
12. Roll, Y., W.D. Cook and B. Golany (1991), Controlling factor weights in data envelopment analysis, *IIE Transactions* 23, 2-9.
13. Rouse, P., M. Putterill and D. Ryan, (1997), Towards a general managerial framework for performance measurement: A comprehensive highway maintenance application, *Journal of Productivity Analysis* 8, 127-149.
14. Rousseau, J., and J. Semple (1993), Categorical outputs in data envelopment analysis, *Management Science* 39, 384–386.
15. Shereman, H.D., and G. Ladino, (1995), Managing bank productivity using data envelopment analysis (DEA), *Interfaces* 25, 60–73.
16. Zhu, J. (2003), Imprecise data envelopment analysis (IDEA): A review and improvement with an application, *European Journal of Operational Research* 144, 513-529
17. Zhu, J. (2004), Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company, *Operations Research* 52, 323-329

This chapter is based upon W.D. Cook, Y. Roll and A. Kazakov, 1990, A DEA Model for Measuring the Relative Efficiency of Highway Maintenance Patrols,INFOR, 28, No. 2, 113-124, and *W.D.* Cook. and J. Zhu, 2001, Output Deterioration with Input Reduction in Data Envelopment Analysis, *IIE Transactions*, 35, 309-320, with permission from INFOR and Taylor & Francis, respectively.

http://www.springer.com/978-0-387-24137-1

Modeling Performance Measurement
Applications and Implementation Issues in DEA
Cook, W.D.; Zhu, J.
2005, XIII, 408 p. With CD-ROM., Hardcover
ISBN: 978-0-387-24137-1