IMPROVING DISCERNMENT IN D.E.A. USING PROFILING

Dr. Christopher Tofallis

The Business School

University of Hertfordshire

College Lane, Hatfield

Hertfordshire

AL10 9AF

Tel. 01707 285 486

Fax. 01707 285 489

e-mail c.tofallis@herts.ac.uk

ABSTRACT

The failure of data envelopment analysis to provide adequate discrimination between a small number of alternatives is illustrated using a facility location problem. A straightforward modification which finds a desirability (or relative efficiency) score for *each* input variable provides greater discrimination and at the same time discerns the strengths and weaknesses of each alternative.

Key words - DEA, efficiency, multicriteria, performance ratios.

IMPROVING DISCERNMENT IN D.E.A. USING PROFILING

INTRODUCTION

Doyle and Green [3] called for the inclusion of DEA (data envelopment analysis) in the set of tools used in multiple criteria decision making after Stewart [7] failed to include it in his survey. Stewart responded in [9] and has since been active in working at the interface between DEA and multi-criteria decision making [8]. Whilst DEA has generated a good deal of attention it does have difficulties associated with it. Stewart's stated aim was to 'identify pitfalls in the usage of various approaches' and we intend to do the same below for DEA, using the same illustrative example employed by all of these authors.

EXAMPLE PROBLEM

The problem is to choose a location for a power plant from among six alternatives by considering six criteria:

z_1 = manpower required

z_2 = power generated in megawatts

z_3 = construction costs in millions of dollars

z_4 = annual maintenance costs in millions of dollars

z_5 = number of villages to be evacuated

z_6 = safety level

High values of safety and power generated are preferred and low values for the other criteria. In DEA the former variables are termed outputs and the latter are the inputs. The desirability D_k of alternative k is expressed as the ratio of the sum of its weighted outputs to the sum of its weighted inputs. (In DEA this quantity is called the relative technical efficiency.) A distinctive feature of DEA is that each alternative is allowed to employ its own set of weights so as to maximise its desirability score subject to the conditions that these weights are positive and that they do not lead to any alternatives having a score exceeding 1 (or 100%). This leads to a linear programming problem for each alternative to discover its optimal set of weights and its resulting score. Of course the freedom to choose optimal weights leads to extremely biased and unrealistic values although work has been done to limit this flexibility by including additional restrictions, often based on the decision maker's beliefs or experience e.g. Dyson and Thanassoulis [5], Roll and Golany [6], Wong and Beasley [11], for a recent survey see Allen et al [1] . A second criticism and one which does not appear to be much appreciated, is also related to this 'weight flexibility' aspect, this is the possibility of

obtaining nonsense ratios. Consider a location which placed all of its weight on criteria 2 and 5, the desirability score then reduces to the ratio: number of megawatts generated per village evacuated. This is probably a meaningless ratio since there is unlikely to be any link between these two quantities. This points out one of the concerns one must have with a method which indiscriminately throws all variables into a single evaluation formula.

The data that was used is reproduced in Table 1. When conventional DEA was applied to it as described above, each alternative was found to have the maximum score of 100% and so no progress was made in discriminating between the alternative locations. Users of DEA are well aware that they must have a large number of alternatives in order for the method to pick out the dominated ones which lie behind (or are 'enveloped' by) the 'efficient frontier'. The more criteria one has the more alternatives are needed - some users look for reasons to discard some criteria from the analysis after discrimination is found to be poor. In this example we have six criteria so the frontier would be a 5-dimensional surface made up of hyperplane segments in a 6-dimensional space. Since there are only six alternative locations we have only six points in this space and in this case they all end up defining this surface with none of them lying behind it because none of them are dominated. The strength of DEA arises when there are dominated alternatives because these can immediately be removed from further consideration in the selection process; clearly if you are free to employ your own (optimal) weights and somebody else uses them to beat you at your own game then a strong statement is being made.

It was then suggested that to improve discrimination the weights used by each location be applied to every other location to obtain what are called cross-evaluation scores and hence an average score could be found for each location based on its own weights as well as those used by the other locations. Due to the fact that DEA always produces non-unique solutions when an alternative has a 100% score, it was necessary to introduce a secondary objective to select from amongst the multiple optima. This employed an 'aggressive' formulation which tended

to reduce the scores of the other locations. This makes one uncomfortable because it again perpetuates the use of extreme and unrealistic weights. To say that this must be fair because every location has the same opportunity to do this ignores the fact that the selection process inadvertently turns into a 'ganging together' of like alternatives: clearly if two locations have similar attribute levels they will employ similar weights and effectively raise each other's average score when these weights are applied to the rest of the sample. If none of the remaining four locations have similar attribute levels they will be disadvantaged because they will be isolated in the cross-evaluation phase. Hence one or both of the two similar locations may turn out to be the winner simply because they effectively gave 'high votes' to each other. Whilst this does not appear to have occurred in an obvious way with our example it does highlight the fact that if an alternative's attributes are very different from the rest of the sample then it stands a much lower chance of winning.

PROFILING

After any multiple criteria decision we ought to be able to explain to others the strengths of the winner in comparison to the others that were not chosen. DEA can be adapted very easily to allow us to do this but in order to do so we will have to move away from the notion of a single score which attempts (but fails) to distil all the data. Instead we may assess the efficiency with which *each* input is being utilised by *relevant* outputs. As the original source of this example [2] lacks the necessary detail I will have to make the following assumptions: that the input variable 'villages evacuated' is related to the safety level but not to power output, and that each of the other inputs (manpower, construction cost, and maintenance cost) is related to both outputs (power generated and safety level). Whether these particular connections are realistic for this illustrative example does not bear upon the general approach being presented. We now use the DEA method but taking *one* input at a time and only with its

related outputs. For example to assess the relative efficiency with which manpower (z_1) is being utilised by alternative *k* solve the following linear programme:

$$\max D_{k1} = (w_{k2} z_{k2} + w_{k6} z_{k6}) / z_{k1}$$

$$\text{subject to } w_{ki} > 0 \text{ for all criteria i}$$

$$\text{and } D_{j1} \leq 100\% \text{ for all sites j}$$

The results of this approach are displayed in Table 2 and can be generated using any DEA software (e.g. *iDEA* produced by Doyle and Green at the University of Bath, or the Warwick University package), or of course by any LP solver. In contrast to the conventional DEA approach we now see a good deal of discrimination between the alternatives. We are also able to discern the strong and weak aspects of each alternative e.g. Belgium is making efficient use of its building expenditure but not its maintenance expenditure. The table shows that Portugal has a higher score than France and Italy in each of the four areas. The fact that France and Italy are each dominated by Portugal is something which conventional DEA did not show. Whilst there is no alternative that scores 100% on all four inputs, Portugal comes top for two of them and has the second highest score for the remaining two. If one used a rule that said the chosen alternative must at least achieve some minimum score, say 50%, on each of these input utilisation scores then Portugal would win. The approach presented here gives a profile for each alternative and hence further assists in 'extracting the most from the data' - a virtue which Stewart picked up on when comparing DEA with mainstream multi-criteria methods [9]. The latter tend to make more use of value judgments. Profiling as presented here may not lead to a clear winner and value judgments may still be necessary in the end, but the above analysis will have extracted more useful information from the data which in turn will assist the decision maker to make a more informed and defensible decision. An application of this approach to data generated from a production function appears in [10].

By taking one input at a time for analysis we have assumed that none of the inputs are substitutes for each other - if two inputs were substitutes and one location made great use of one and little use of the other then it would find it had a low efficiency score for the former and a high score for the latter. Clearly in such a case one needs to take a linear combination of the two inputs and let the linear programme find the relevant weights. Only if every input variable can act as a substitute for every other should we include them all together in the expression for desirability (D_k), we would then end up with the conventional DEA formulation; sadly, it has invariably not been recognised that this assumption is being made when the technique has been applied in the past.

One might ask why one should not proceed to carry out a profiling exercise for the outputs too i.e. to analyse them separately, taking the desirability score as the output divided by a weighted combination of related inputs. (Doing this shows only Portugal and the U.K. to have scores of 100% for each output.) However, whereas a given input may be 'used up' or shared among the relevant outputs, it is not true that the full amount of each of the relevant inputs will be consumed by a single output, so taking the ratio in the latter case would be difficult to justify.

CONCLUDING REMARKS

This memorandum calls for a more judicious use of DEA, firstly by assessing together only those variables (inputs and outputs) which are linked, and secondly by assessing separately the relative efficiency with which *each* input is being utilised. Doing so provides greater discrimination and also show the weaknesses and strengths of each alternative. Both of these consequences will aid the decision maker in making a more informed decision which can also be better justified to others.

Like Stewart I wish to thank Doyle and Green for bringing DEA to the attention of a wider audience. It has the power to screen out dominated alternatives and, as we have shown, its potential for extracting useful information from the data has not been fully explored.

REFERENCES

1. Allen R , Athanassopoulos A, Dyson RG, Thanassoulis E (1994) Weights restrictions and value judgments in DEA. Warwick Business School Research Paper 138.

2. Brans JP and Mareschal B (1990) The Promethee methods for MCDM. In *Readings in Multiple Criteria Decision Aid.* Ed. Bana e Costa CA . Springer-Verlag.

3. Doyle J and Green RH (1993). Data envelopment analysis and multiple criteria decision making. *Omega* **21**, 713-715.

4. Doyle J and Green RH(1989). iDEA (Integrated DEA) software, version 1.2.1. School of Management, University of Bath.

5. Dyson RG and Thanassoulis E (1988). Reducing weight flexibility in DEA. *J. Opl Res. Soc*. **39**, 563-576.

6. Roll Y and Golany B (1993). Alternate methods of treating factor weights in DEA. *Omega* **21**, 99-109.

7. Stewart TJ (1992). A critical survey on the status of multiple criteria decision making theory and practice. *Omega* **20**, 569-586.

8. Stewart TJ (1994) Relationships between data envelopment analysis and multicriteria decision analysis.University of Cape Town working paper.

9. Stewart TJ (1994). Data envelopment analysis and multiple criteria decision making: a response. *Omega* **22**, 205-206.

10. Tofallis C (1994) Resource efficiency profiling. University of Hertfordshire Business School Working Paper. Also to appear in *Multi-Objective and Goal Programming: Theories and Applications* (1995) ed. Tamiz M. Springer-Verlag.

11. Wong Y-HB and Beasley JE (1990). Restricting weight flexibility in DEA. *J. Opl Res. Soc.* **41**, 829-835.

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
ITALY	80	90	600	54	8	5
BELGIUM	65	58	200	97	1	1
GERMANY	83	60	400	72	4	7
UK	40	80	1000	75	7	10
PORTUGAL	52	72	600	20	3	8
FRANCE	94	96	700	36	5	6

TABLE 1. Attribute values for six locations.

	CONSTRUCTION COST	MANPOWER	MAINTENANCE COST	VILLAGES EVACUATED
ITALY	70	56	46	23
BELGIUM	100	45	17	37.5
GERMANY	100	36	24	66
UK	57	100	33	54
PORTUGAL	78	69	100	100
FRANCE	67	51	74	45

TABLE 2. Profiles showing the percentage efficiency with which each input is being utilised.