A Cautionary Note on Methods of Comparing Programmatic Efficiency Between two or more Groups of DMUs in Data Envelopment Analysis.

Dr Gary Simpson

Aston Business School

Aston University

Birmingham B4 7ET

Tel +44 (0) 121 204 3241

Fax +44 (0) 121 204 5271

Email: g.p.m.simpson@aston.ac.uk

Abstract

In some applications of Data Envelopment Analysis (DEA) there may be doubt as to whether all the DMUs form a single group with a common efficiency distribution. The Mann-Whitney rank statistic has been used to evaluate if two groups of DMUs come from a common efficiency distribution under the assumption of them sharing a common frontier and to test if the two groups have a common frontier. These procedures have subsequently been extended using the Kruskal-Wallis rank statistic to consider more than two groups. This technical note identifies problems with the second of these applications of both the Mann-Whitney and Kruskal-Wallis rank statistics. It also considers possible alternative methods of testing if groups have a common frontier, and the difficulties of disaggregating managerial and programmatic efficiency within a non-parametric framework.

Keywords: Data Envelopment Analysis (DEA); Statistics, Programmatic Efficiency.

In one of the first empirical applications of the DEA methodology Charnes, Cooper and Rhodes (1981) introduce a distinction between two types of efficiency which are of interest to policymakers. These two types of efficiency are managerial efficiency and programmatic efficiency. The managerial efficiency is the classic DEA efficiency and measures the performance of an individual decision making unit (DMU) in comparison with the observed production possibility frontier. The programmatic efficiency recognises that different groups of the DMUs may not have the same production possibility frontier because of programmatic differences and seeks to reveal potential efficiency differences between the productive programmes and to test the relative efficiency of each programme irrespective of potentially different distributions of managerial efficiency between them.

Brockett and Golany (1996) re-analysed the data of Charnes, Cooper and Rhodes (1981) and proposed a process using the Mann-Whitney rank statistic to test if a group of DMUs representing one program is more efficient than another by the nature of the program as opposed to the efficiencies of the individual DMUs within it. These ideas have been subsequently extended to more than two groups by using a more general rank sum test, the Kruskal-Wallis rank test (Sueyoshi T, Aoki S (2001)). Both papers have been quite widely referenced, some citations such as Revilla E et al (2003) merely mention the papers as others that have used either the Mann-Whitney or Kruskal-Wallis test statistics in DEA studies before correctly applying these within their own study but other citations including Brockett et al (2005), Ross AD and Droge C (2004), Golany B, Storbeck JE (1999) also apply the techniques described in the papers to test for programmatic changes in the frontier. The technique is also described as a standard technique in Cooper, Seiford and Tone's (2000) popular DEA text book.

The following steps are proposed by Brockett and Golany (1996) to estimate the programmatic efficiency (as opposed to managerial efficiency).

1. Split the DMUs into the two groups according to their respective programmes and run DEA separately for each one of the two groups;

2. In each of the two groups project the inefficient DMUs to the efficient frontier of this group, thus attempting to eliminate the effects of managerial inefficiencies within a programme.

3. Run DEA on the combined set of the projected DMUs from both groups; which includes all DMUs.

4. Apply a statistical test to test if the two groups have the same distribution of efficiency values.

First we will describe the difficulties in separating programmatic and managerial inefficiency within a DEA framework, and then we will demonstrate that the process used by Brockett and Golany is inappropriate and will invalidate the results of the statistical test. This is because of the process at step 2 above which will depend on how the efficient DMUs are distributed between the programmes. It is also demonstrated that it produces biased results, particularly favouring the larger of two unequally sized programmes. We then show that these problems continue when more than two groups are considered and present some simulation results. Finally we make some suggestions as to how these problems maybe addressed.

1.   *Managerial Efficiency vs Programmatic Efficiency*

If different programmes transform the same inputs into the same outputs there is said to be no programmatic inefficiency if they share a common production possibility frontier. That is the outputs that can be obtained for a given set of inputs are the same for an efficient DMU in each of the programmes. Managerial inefficiency refers to any short fall of the outputs or over consumption of inputs of a particular DMU relative to the production possibility frontier.

So considering Figure 1, if the DMU A shown is a member of Programme A its managerial efficiency is given by OA/OA' and the programmatic efficiency of Programme A for DMUs with this output mix would be given by OA'/OA''. Whereas, if DMU A shown is a member of Programme B its managerial efficiency is given by OA/OA'' and the programmatic efficiency of Programme A for DMUs with this output mix would be given by OA''/OA''=1. Hence, this DMU provides no evidence of programmatic inefficiency for programme B but does provide evidence of programmatic inefficiency for programme A.

In Data Envelopment Analysis the observed best practice frontier is used rather than the true but unknown production possibility frontier. Because, the observed frontiers of two randomly selected groups of DMUs having a common production possibility frontier would be expected to differ, there is a desire to ascertain if the observed differences provide sufficient evidence at a given level of significance to reject the null hypothesis that the production possibility frontiers are the same.

Testing this hypothesis is complicated by the expectation that the managerial efficiencies of the programmes will differ. For example, if one of the programmes is a new initiative which was thought

to have attracted better managers then different distributions of managerial efficiency would result; and so it is desirable to allow for different distributions of managerial efficiency. It is such a process that Brockett and Golany proposed; correcting for managerial inefficiency by projecting to the within programme frontier. Unfortunately because this projection step is not equally effective for each of the programmes at removing the managerial inefficiency, their test will generally be biased against programmes represented by fewer DMUs and/or with programmes with more managerial inefficiency. We can at this point also note that there is an intrinsic difficulty in disaggregating managerial and programmatic efficiency. This is because if we do not impose any restrictions of the distributions of managerial efficiency for the groups any feasible differences can be entirely explained by the differing distributions of managerial efficiency in the groups. For example consider the case where a new programme has selected the best 5% of managers overall, even if there is no difference in the true production frontiers this new programme will dominate the observed frontier unless more than 5% of managers are 100% efficient! Similarly in a case where one programme has an advantage this may not be observed if its managers are less efficient. While one programme may have a reasonable proportion of its managers on, or very close to the true frontier, this may not be the case for the other programme as it may have lost many, and possibly all, of its best managers. That is, in principle we should not expect projection to the within programme frontier to correct for managerial efficiency because such projections are both dependent on the sample size and the distributions of managerial efficiency of the programmes.

2.Problems with Brockett & Golany's programme evaluation procedure

The main difficulties with Brockett and Golany's proposed method is the correction of managerial inefficiency by projecting to the within programme frontier and the subsequent use of projected values

in the Mann-Whitney test. This is problematic for two reasons; firstly the projection to the within programme frontier is unlikely to be equally effective at correcting for managerial efficiency particularly for unequally sized programmes, and secondly the efficiencies of the projected DMUs are not independent but are determined by the comparative efficiencies of observed frontiers of the two programmes. We will initially demonstrate that this procedure results in seriously erroneous results by considering a simple case.

Let us assume that we have 100 DMUs which all share a common production possibility frontier and are drawn from a common managerial efficiency distribution. We divide the set of DMUs into two groups at random. Hence we know there is no difference in programmatic efficiency between the two groups.

We then apply the procedure proposed as follows:

    I.   Split the group of all DMUs into two groups (A and B) containing n_a and n_b DMUs respectively. Run DEA separately for each of the two groups.

    II.   In each of the two groups separately, adjust each inefficient DMU to its "level if efficient" value by projecting it onto the efficiency frontier for that group.

    III.   Run a pooled DEA with all the projected DMUs

    IV.   Apply Mann-Whitney Rank Test to the pooled DEA results and compute

$$Z = \frac{\frac{n_a}{2}(n_a + n_b + 1) - \sum_{GroupA} Ranks}{\sqrt{\frac{n_a \cdot n_b (n_a + n_b + 1)}{12}}}$$

Let us assume initially that the two randomly selected groups are Group A consisting of 24 DMUs and Group B consisting of the remaining 76 DMUs and that 10 of these 76 DMUs define the efficient frontier of the whole set of 100 DMUs.

The probability that all 10 of these globally efficient DMUs happen to have been put in Group B by chance is given by

Probability $= {}^{76}C_{10} / {}^{100}C_{10} = 0.0551$.

Now, let us project DMUs in each of Groups A and B to their own respective frontiers and then recalculate the efficiencies of the pooled DMUs.

All 76 DMUs in Group B will appear efficient as all ten globally efficient DMUs appear in this group and so also define its frontier. Hence all 66 of the inefficient DMUs in Group B will be projected to this globally efficient frontier.

However none of the DMUs in Group A will appear on the joint efficient frontier. So none of the projected DMUs in this group will be projected to the joint frontier.

Hence, the DMUs in Group B will occupy ranks 1 to 76, and have a Rank Sum = 2926.

Whereas the DMUs in Group A will occupy ranks 77 to 100 and have a Rank Sum = 2124.

So calculating the Mann-Whitney Rank Statistic gives Z=-7.36 suggesting there is almost no chance that the null hypothesis is true (P=0.0000).

That is the test statistic suggests that this distribution is highly significant and there is practically no chance of the groups sharing a common frontier but it will occur in greater than 5% of cases when there is no difference between the two groups either in terms of the production frontier or managerial efficiencies! This problem will persist when the groups are of more similar sizes and when not all of the efficient DMUs fall in one of the groups, as we will illustrate in the following example.

Again consider 100 DMUs but now divided into two randomly selected groups; Group A consisting of 37 DMUs and Group B consisting of the remaining 63 DMUs and again 10 of the DMUs define the observed efficient frontier of the whole set of 100 DMUs.

The probability that all 10 of these globally efficient DMUs happen to have been put in Group B by chance is now given by

Probability $= {}^{63}C_{10} / {}^{100}C_{10} = 0.0074$

which is quite small, but still considerably larger that the p-value that would be obtained by applying the procedure of Brocket and Golany. Let us now consider the probability that 9 of these globally efficient DMUs happen to have been put in Group B and 1 is placed in group A.

Probability $= {}^{63}C_9 \, {}^{37}C_1 / {}^{100}C_{10} = 0.0506$.

We now consider what happens when we project the DMUs to their respective frontiers. For the sake of simplicity we will consider a one-input two-output constant returns to scale DEA model as shown in Figure 2. We will make no assumption about the distribution of the inefficient DMUs but will continue with the assumption that the globally efficient DMUs are randomly distributed between the two groups. Let η_i be the number of inefficient DMUs in group B that project onto the ith segment of the frontier (numbering the segments in order of increasing ratio of output2 to output1). Then, for segments 2 to 10 (that is all the segments on the non-dominated part of the frontier) the inefficient DMUs in Group B will project to the joint efficient frontier if and only if both of the DMUs defining the segment are in Group B. The probability of both being in group B is 8/10, but for segments 1 and 11 only a single DMU defines the segment so the probability becomes 9/10.

Hence we expect $9 + \dfrac{8}{10} \sum\limits_{i=2}^{10} \eta_i + \dfrac{9}{10}(\eta_1 + \eta_{11}) \geq 9 + \dfrac{8}{10} \sum\limits_{i=1}^{11} \eta_i = 9 + \dfrac{8}{10}(63 - 9) = 52.2$ of the DMUs in

Group B to appear on to the joint efficient frontier. That is the nine DMUs which defined part of the

frontier and more than 80% of the inefficient DMUs in group B are expected to appear on the joint frontier.

 Whereas inefficient DMUs from Group A will only project to the joint frontier if they are un-enveloped and the efficient DMU in Group A is the closest one as shown in figure 2.

So we expect $1+\dfrac{0}{10}\sum\limits_{i=2}^{10}\eta_i+\dfrac{1}{10}(\eta_1+\eta_{11})\leq 1+\dfrac{1}{10}\sum\limits_{i=1}^{11}\eta_i=1+\dfrac{1}{10}(37-1)=4.6$ DMUs from Group A to appear on the joint efficient frontier.

That is a total of about 57 DMUs are expected to appear on the joint efficient frontier and to tie for the top rank and will each be given a rank number of 29. The remaining 43 DMUs will share the remaining ranks (58 to 100) with an average rank number of 79.

So the expected rank sum for group A is 4.6x29+32.4x79 =2693

This would give Z =-5.89 and again P=0.0000. So once again the test has suggested a highly significant difference when the result is in fact to be expected in more than 5% of the cases. We will demonstrate that these problems persist even with equal group sizes by means of a simulation study, but first we consider what happens when there are more than two groups.

3 More than two groups: the Kruskal and Wallis Rank Test

As already stated Sueyoshi & Aoki extend the ideas of Brockett & Golany in line with the suggested extensions in their original paper to consider more than two groups of DMUs and use it to consider how a frontier may shift over time. Their procedure works in the same way projecting DMUs to the within group frontier then pooling the projected DMUs and ranking them using DEA. But now the sums of the ranks are compared using the Kruskal Wallis Rank test.

If there are K groups and the jth group has n_j DMUs and Rank Sums R_j and $N=\Sigma n_j$

then $H = \left(\dfrac{12}{N(N+1)} \right) \sum_{j=1}^{K} \dfrac{R_j^2}{n_j} - 3(N+1)$ should follow a χ^2 distribution with *K-1* degrees of freedom.

We will again demonstrate that this procedure will be problematic by way of an illustrative example. We take a set of 60 DMUs which all share a common production possibility frontier and are drawn from a common managerial efficiency distribution, and divide the set of DMUs into three groups at random. Hence we know there is no difference in programmatic efficiency between the three groups. We will now consider each of the three randomly selected groups to consist of 20 DMUs, as have all the groups the same size should minimise the problems with the procedure, and that 5 of the DMUs define the observed efficient frontier of the whole set of 60 DMUs.

Firstly we note that the probability that all 5 globally efficient DMUs are in one of the groups is given by : Probability $= 3 \, {}^{20}C_5 / {}^{60}C_5 = 0.00852$.

The group which possesses all the globally efficient DMUs would take up ranks 1 to 19 and the remaining ranks would be shared by the other two groups. Hence we will have a minimum value of H when the remaining ranks are equally shared between these two groups of

$$H = \left(\frac{12}{60(60+1)} \right) \left(\frac{210^2 + 810^2 + 810^2}{20} \right) - 3(60+1) = 39.34 \,.$$

This would in turn give a p-value of 0.00000. So the probability of rejecting a true null hypothesis is again considerably larger than the test suggests.

We now consider a less extreme distribution of the globally efficient DMUS with 4 of the globally efficient DMUs are in one of the groups, 1 of them in the second and none in the third group. Such a distribution is to be expected in more than 10% of the cases if the groups were randomly selected (Probability $= 6 \, {}^{20}C_4 \, {}^{20}C_1 / {}^{60}C_5 = 0.10645$). This is illustrated in Figure 4, and we will demonstrate the problems persist using the same logic as we did for the two group case.

For Group A (the group with none of the globally efficient units) clearly none of its inefficient DMUs will project to the joint frontier. With Group B (with only one of the efficient units) the expected number of its inefficient DMUs that will project to the joint frontier is given by

$$\frac{0}{5} \sum_{i=2}^{5} \eta_i + \frac{1}{5}\left(\eta_1 + \eta_6\right) \le \frac{1}{5} \sum_{i=1}^{6} \eta_i = \frac{1}{5}\left(20 - 1\right) = 3.8.$$

So in total we expect less than 4.8 DMUs from this group on the joint frontier (3.8 plus the 1 globally efficient unit.). But for Group C (the first group with 4 of the globally efficient DMUs) the expected number of its inefficient DMUs that will project to the joint frontier is given by

$$\frac{3}{5} \sum_{i=2}^{5} \eta_i + \frac{4}{5}\left(\eta_1 + \eta_6\right) \ge \frac{3}{5} \sum_{i=1}^{6} \eta_i = \frac{3}{5}\left(20 - 4\right) = 9.6.$$

So in total we expect at least 4+9.6=13.6 of the DMUs from this group on the joint frontier. Hence, for all the groups we expect a total of 18.4 DMUs on the joint frontier (hence an average rank of 9.7) with the remaining 41.6 DMUs sharing the remaining ranks (with an average rank of 39.7). This gives the following expected rank sums, under the conservative assumption that the remaining ranks are shared equally between the programmes.

Group A Rank Sum = 20x39.7 =794

Group B Rank Sum = 4.8x9.7 +15.2x39.7 = 650

Group C Rank Sum = 13.6x9.7+6.4x39.7 = 386

Then the expected value of H is greater than

$$H = \left(\frac{12}{60(60+1)}\right)\left(\frac{386^2 + 650^2 + 794^2}{20}\right) - 3(60+1) = 14.0327$$

which would give a p-value of less than 0.000895.

So again the procedure has produced a highly significant result, but we expect 4 of the 5 globally efficient DMUs to be in the same group more than 10% of the time if the DMUS are randomly assigned to the Groups! In this case, there is not bias against any particular programme, as each, being the same size are equally likely to be under-represented in the number of DMUs on the joint frontier prior to projection. There is a bias however against any programme which are under represented in this way because the projection of the inefficient DMUs to the within group frontier exaggerates the differences between the groups and results in the over sensitivity of the test. Hence the test suggests a programmatic difference where none exists.

In the following section we present simulation results that demonstrate that this problem persists for larger groups.

4 Simulation Results

In the preceding section we have attempted to demonstrate the problems that exist with the existing tests of programmatic differences at a conceptual level. We now present the results from a simulation study further highlighting these problems. We generate 6000 DMUs using the same production function and independently draw managerial efficiency for each DMU from a half normal distribution. We divide the DMUs into 80 groups of 25 DMUs, 40 groups of 50 DMUs and 20 groups of 100 DMUs and apply Brockett and Golany's procedure to different pairs of groups. We will also consider some results for three or more groups.

The production function chosen is of the Cobb-Douglas form $y = 0.75x_1^{0.65} x_2^{0.55}$ and is essentially the same as the one used by Bardhan, Cooper and Kumbhakar, (1998) with the exception that we do not include a normally distributed measurement error. Similarly x_1 is drawn from a uniform distribution over the interval [500,1000] and x_2 is drawn from a uniform distribution over the interval [400,600].

We now add inefficiency by calculating the observed inputs $\hat{x}_1 = x_1 e^{\tau 1}$ and $\hat{x}_2 = x_2 e^{\tau 2}$ where τ_1 and τ_2 represent the input specific technical inefficiencies for each of the inputs and are independently drawn from half-normal distributions with mean 0.12 and standard deviation 0.1.

Table 1 below shows the percentage of times Brockett and Golany's procedure finds the difference between two groups generated in this way significant at the 1%, 5% and 10% levels of significance. Clearly these should always be approximately 1%, 5% and 10% respectively if the test is working correctly.

Compared group sizes	Number of tests	level of significance		
		1%	5%	10%
100 DMUs	190	53%	65%	69%
50 DMUs	374	60%	70%	76%
25 DMUs	385	32%	46%	54%

Table 1: Comparing two equally sized groups.

It can be clearly seen that even for large equally sized groups Brockett and Golany's procedure will generate an unacceptably high level of false positives. In fact there is no evidence that increasing the group size improves the performance of the test. This is because although the probability of the most extreme result decreases with increasing sample size, this is balanced by the lack of independence in the efficiencies of the projected DMUs which becomes more problematic as sample size increases. The situation is even worse when we consider unequally sized groups; see Table 2 below. The larger group is almost invariably identified as having a superior frontier, so the statistic not only generates an unacceptable level of false positives but is strongly biased in favour of the larger group.

Compared group sizes		Number of tests	Group A superior level of significance		Group B superior level of significance	
Group A	Group B		5%	10%	5%	10%
100 DMUs	25 DMUs	100	90%	92%	0%	1%
100 DMUs	50 DMUs	100	80%	84%	1%	1%
50 DMUs	25 DMUs	100	62%	66%	5%	6%

Table 2: Comparing two unequally sized groups.

Table 3 reports similarly extreme results for comparing 3 equally sized groups using the Kruskal Wallis Rank test suggested by Sueyoshi & Aoki.

Compared group sizes	Number of tests	level of significance		
		1%	5%	10%
100 DMUs	125	50.4%	64.0%	72.8%
50 DMUs	125	66.4%	76.0%	79.2%
25 DMUs	125	51.2%	61.6%	68.0%

Table 3: Percentage of times null hypothesis rejected at different levels of significance for comparing three equally sized groups.

In summary, we conclude that current tests for programmatic efficiency are seriously flawed. They have strong tendencies to find programmatic differences where none exists and a clear bias towards larger programmes. We have demonstrated why this should be expected from a theoretical point of view and demonstrated the extent of the problems with a simulation study.

5 Possible Ways Ahead

It is clear that the current tests do not properly separate managerial and programmatic efficiency.

The root cause of the problems is in the projection of units to the within programme frontier, which does not treat the programmes equivalently, particularly when the number of DMUs in the programs are unequal. We suggest that these tests should no longer be presented as a standard method of testing for programmatic efficiency differences in a DEA framework. We do not believe a general solution to the problems we highlight is possible, however if panel data is available or if one is prepared to make assumptions about the distribution of managerial efficiency some progress can be made.

For example, one can note that if rather than projecting the units before applying the non-parametric tests we simply applied them to the un-projected DMUs we have a test that will fairly detect shift in

Programmatic Efficiency Comparisons

the programme frontier if the distribution of managerial efficiency is the same in the two groups. Unfortunately, it is this assumption of a common distribution of managerial efficiency between the programs or groups that we wish to avoid, but if we have data for several time periods this approach may prove fruitful. For example, if we have a single programme into which a new initiative is introduced into a subset of the DMUs, we could perhaps assume the overall distribution of managerial efficiency across all groups is the same before and after the introduction. Applying the Mann-Whitney test without projection to the whole data sets before and after would then test for a shift in the frontier caused by the introduction of the new initiative. A similar test of the subset that introduced the initiative before and after would give the combined effect of the shift in frontier and any advantage in selecting the better managers. If a longer time series is available one could similarly test if the change over time was greater at the point of introduction of the new programme.

Alternatively, if we assume that when there is no programmatic inefficiency there is no association between which facets of the frontier are defined by DMUs in a particular programme and the facets that the inefficient units in that program will project to. That is we expect the input/output mixes to have the same distribution for the two programmes unless one programme has an advantage at a particular mix we could precede with a test on these lines. Simpson (2005) has suggested such a test for the two programme case based on this approach and further work may allow a generalisation to the k programme case. Unfortunately, such tests require an assumption that the input/output mix will only change if the frontiers are different for the two groups, and this is not generally appropriate.

The final possibility and most generally applicable is to make assumptions about either the probability distributions of inefficiency for each of the two groups, and/or the nature of the frontier shift between

the programmes and use a bootstrapping methodology such as that introduced by Simar and Wilson (1998, 2000) but not to our knowledge currently applied in the area of programmatic efficiency comparisons. We believe this is likely to be the most generally applicable solution and call for further research applying bootstrapping techniques to the problem of programmatic efficiency.

References

Banker R.D., (1993), "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation" Management Science, 39, 10 1265-1273

Bardhan, Cooper and Kumbhakar, (1998) "A Simulation Study of Joint Uses of Data Envelopment Analysis and Statistical Regressions for Production Function Estimation and Efficiency Evaluation" Journal of Productivity Analysis, 9, 249-278.

Brockett PL, Golany B, (1996), "Using Rank Statistics for Determining Programmatic Efficiency Differences in Data Envelopment Analysis" Management Science., 42, 3 466-472.

Brockett PL, Cooper WW, Golden LL, Rousseau JJ, Wang YY (2005), "Financial intermediary versus production approach to efficiency of marketing distribution systems and organizational structure of insurance companies", JOURNAL OF RISK AND INSURANCE 72 (3): 393-412

Charnes A, Cooper WW, Rhodes E, (1981), "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through" Management Science, 27, 6 668-697.

Cooper, Seiford and Tone(2000) "Data Envelopment Analysis: a comprehensive text with models, applications, references and DEA-Solver software" ISBN 0-7923-8693-0 page 201.

Golany B, Storbeck JE (1999) "A data envelopment analysis of the operational efficiency of bank branches" INTERFACES 29 (3): 14-26

Revilla E, Sarkis J, Modrego A (2003) "Evaluating performance of public-private research collaborations: A DEA analysis", J Opnl Res Soc 54: 165-174

Ross AD and Droge C (2004) "An analysis of operations efficiency in large-scale distribution systems", JOURNAL OF OPERATIONS MANAGEMENT 21 (6): 673-688

Simar, L and Wilson P W (1998) "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric models" Management Science 44, 49-61.

Simar, L and Wilson P W (2000) "Statistical inference in nonparametric frontier models: the state of the art", Journal of Productivity Analysis 13, 49-78.

Sueyoshi T, Aoki S, (2001), "A use of a nonparametric statistic for DEA frontier shift: the Kruskal and Wallis rank test" Omega Int J Manage Sci, 29 1-18.

Simpson G, (2005), Programmatic Efficiency Comparisons Between Unequally Sized Groups of DMUs in DEA, Journal of the Operational Research Society Vol 56 pages 1431-1438

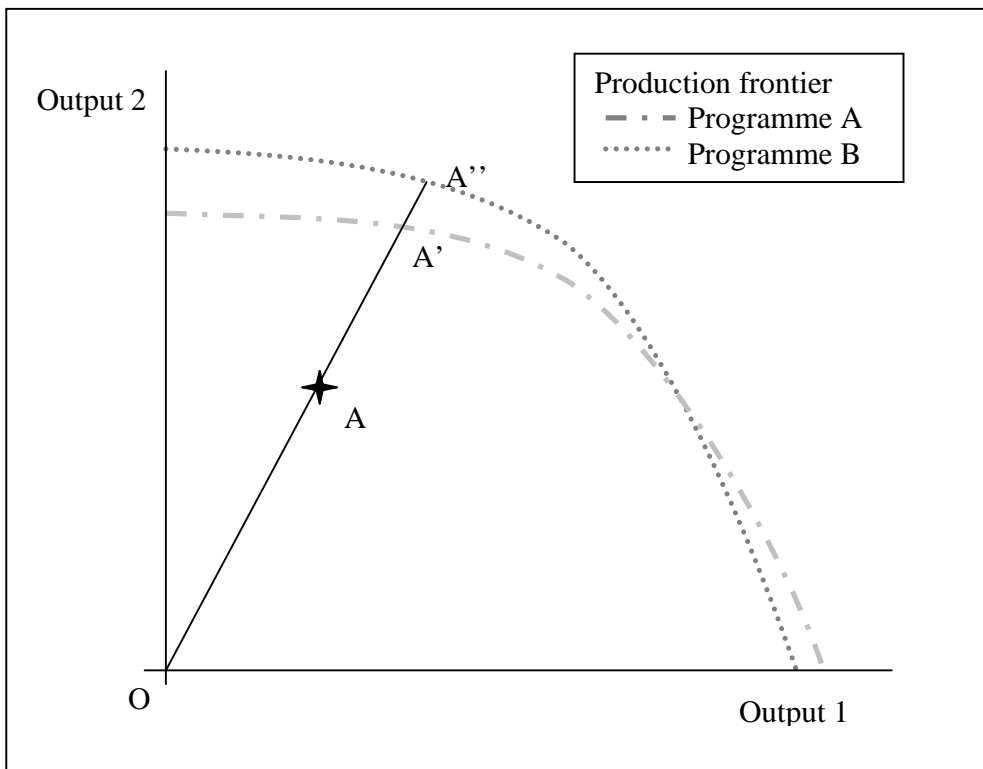Figure 1: Managerial Vs Programmatic Efficiency
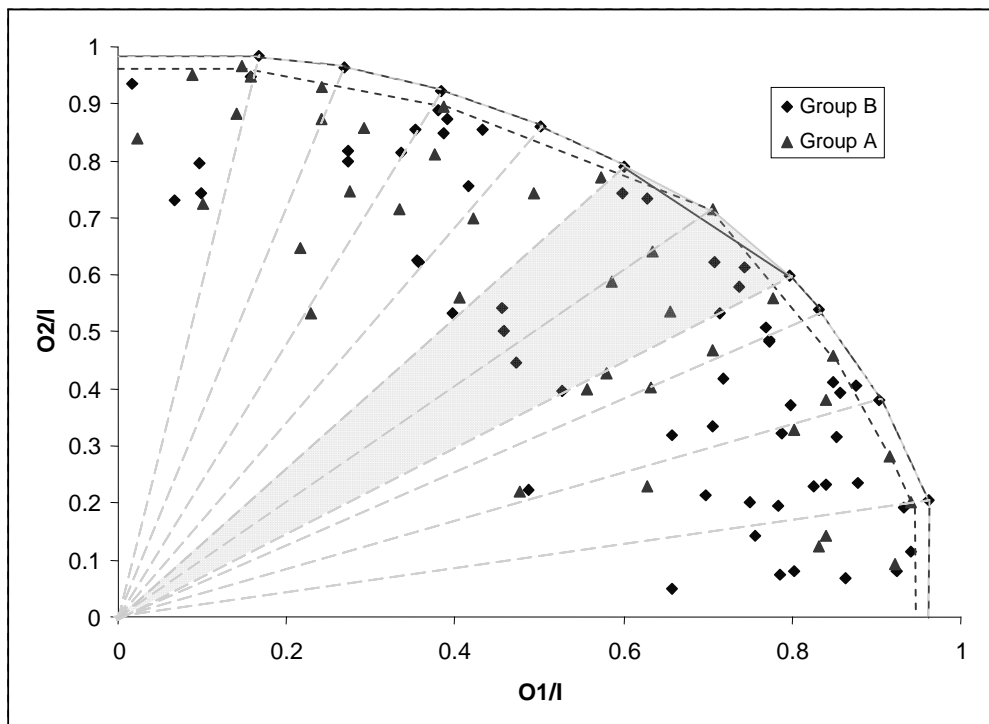

Figure 2: Segments of the DEA frontier

Figure 3: A Segment of the DEA frontier where units from Group A project to the joint frontier
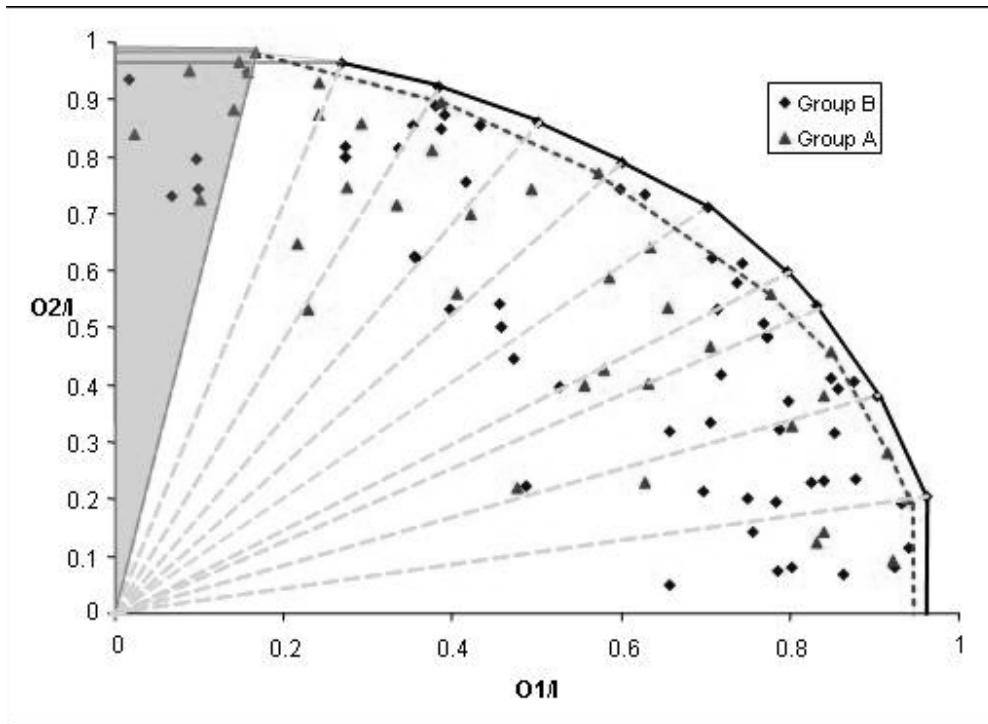

Programmatic Efficiency Comparisons

Figure 4: Production frontiers for 3 groups

Programmatic Efficiency Comparisons