IE 485

DATA ENVELOPMENT ANALYSIS

Data Envelopment Analysis

- A performance measurement technique, which can be used for evaluating the relative efficiency of decision making units (DMUs) in organisations

- DMU: a distinct unit within an organisation that has flexibility with respect to some of the decisions it makes, but not necessarily complete freedom with respect to these decisions

(banks, police stations, hospitals, schools, university departments etc)

A simple example
Graphical Approach

- Consider a number of hospitals. For each hospital we have a single output measure (number of patients served, excluding the ones that need operation) and a single input measure (number of staff).
- The data we have is as follows:

Hospital	Number of patients ('000s)	Number of staff
Croydon	1250	180
Dorking	440	160
Redhill	800	170
Reigate	230	110

How then can we compare these branches and measure their performance using this data?

Ratios

- A commonly used method is *ratios*.
- Typically we take some output measure and divide it by some input measure. Note the terminology here, we view hospitals as taking *inputs* and converting them (with varying degrees of efficiency, as we shall see below) into *outputs*.
- For our hospital branch example we have a single input and a single output measure

Hospital	**Patients served by staff member ('000s)**
Croydon	6.94
Dorking	2.75
Redhill	4.71
Reigate	2.09

Relative Efficiency

- Calculate the *relative efficiency* of other hospitals with respect to A.

Hospital	**Relative efficiency**
Croydon	100%
Dorking	40%
Redhill	68%
Reigate	30%

We could use this comparison with Croydon to set *targets* for the other hospitals.

We could set a target for Reigate of continuing to process the same level of output but with one less member of staff. This is an example of an ***input target*** as it deals with an input measure.

An example of an ***output target*** would be for Reigate to increase the number of patients served by 10% (e.g. by attracting new patients).

We can also set a mix of input and output targets which we want the DMU to achieve.

- Typically we have more than one input and one output.
- Suppose now that we have two output measures
  - number of patients served (without operation)
  - number of operations completed
- Single input measure (number of staff) as before.

Hospital	**Number of patients served ('000s)**	**Number of operations completed**	**Number of staff**
Croydon	1250	500	180
Dorking	440	200	160
Redhill	800	550	170
Reigate	230	120	110

Hospital	Patients served per staff member ('000s)	Operations completed per staff member
Croydon	6.94	2.78
Dorking	2.75	1.25
Redhill	4.71	3.24
Reigate	2.09	1.09

Different ratios can give a different picture and it is difficult to combine the entire set of ratios into a single numeric judgement.

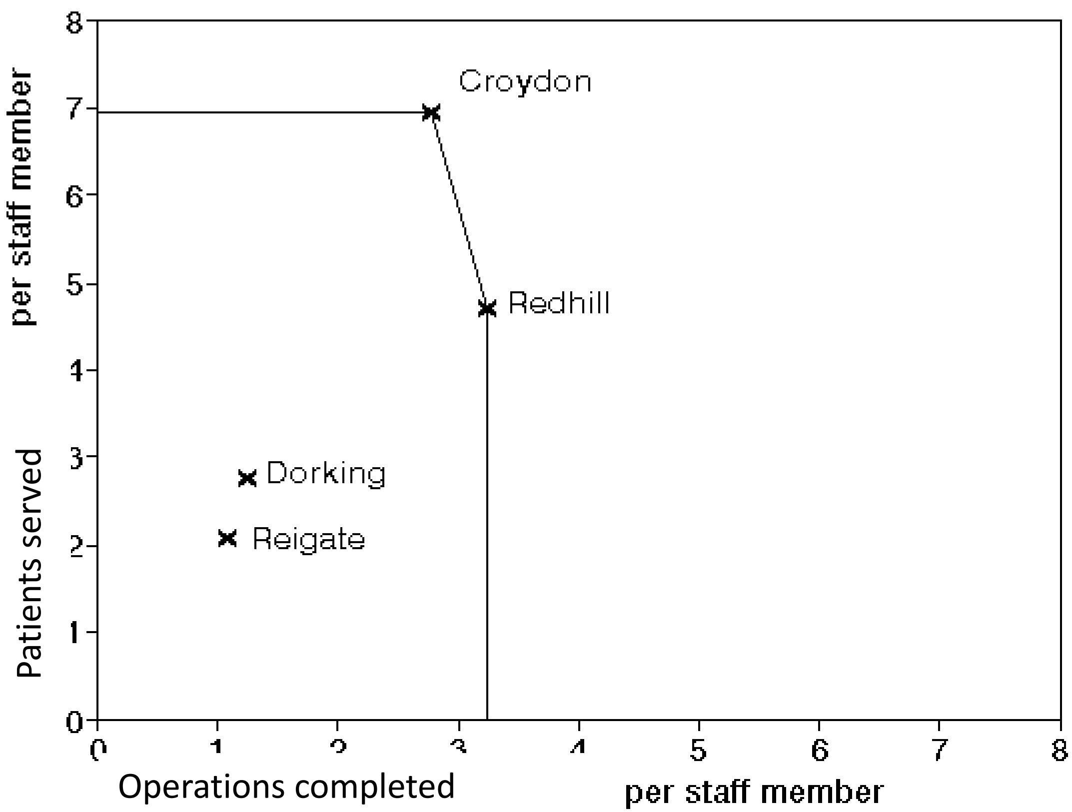
Consider Dorking and Reigate :
Dorking is (2.75/2.09) = 1.32 times as efficient as Reigate at patients served but only (1.25/1.09) = 1.15 times as efficient at operations.

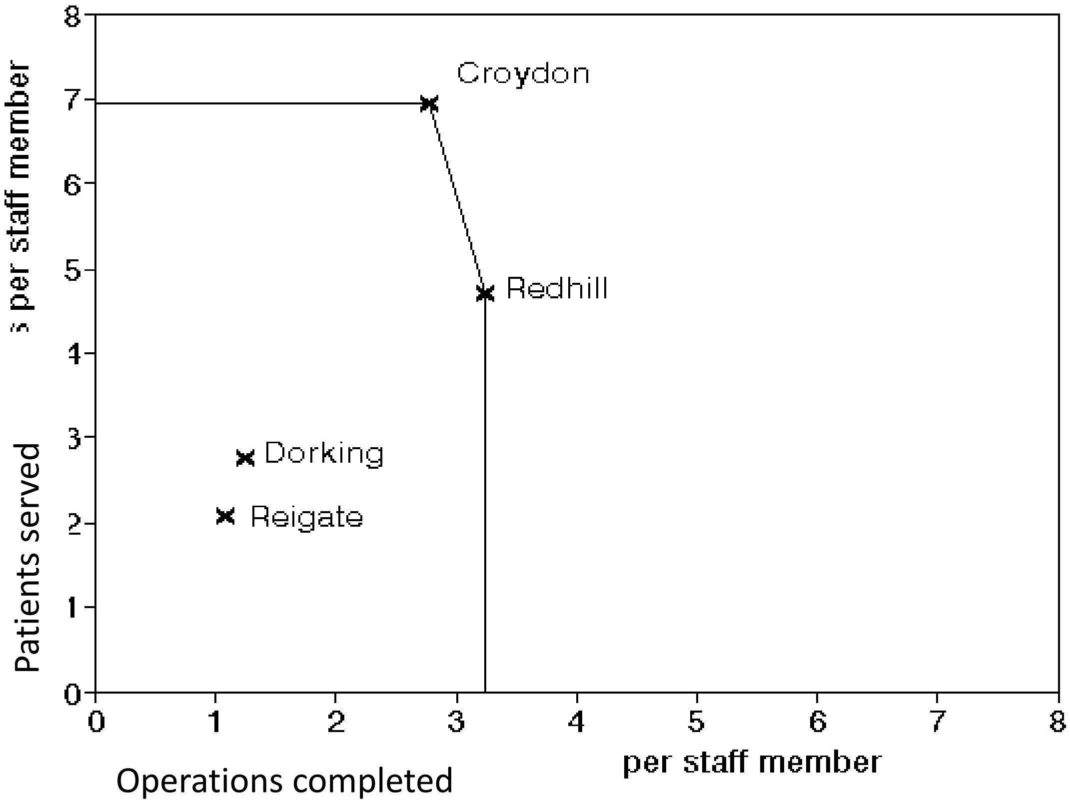
How would you combine these figures into a single judgement?

This problem of different ratios giving different pictures would be especially true if we were to increase the number of branches (and/or increase the number of input/output measures).

- One way around the problem of interpreting different ratios, at least for problems involving just two outputs and a single input, is a simple graphical analysis.

- Suppose we plot the two ratios for each branch





The line is called the *efficient frontier* (sometimes also referred to as the *efficiency frontier*).

Mathematically the efficient frontier is the convex hull of the data.

The efficient frontier, derived from the examples of best practice contained in the data we have considered, represents a standard of performance that the branches not on the efficient frontier could try to achieve.

You can see therefore how the name *data envelopment analysis* arises - the efficient frontier envelopes (encloses) all the data we have.

We say that any branches on the efficient frontier are 100% efficient (have an efficiency of 100%).

This is not to say that the performance of Croydon and/or Redhill could not be improved.

However we can say that, *on the evidence (data) available,* we have no idea of the extent to which their performance can be improved.

Note that:

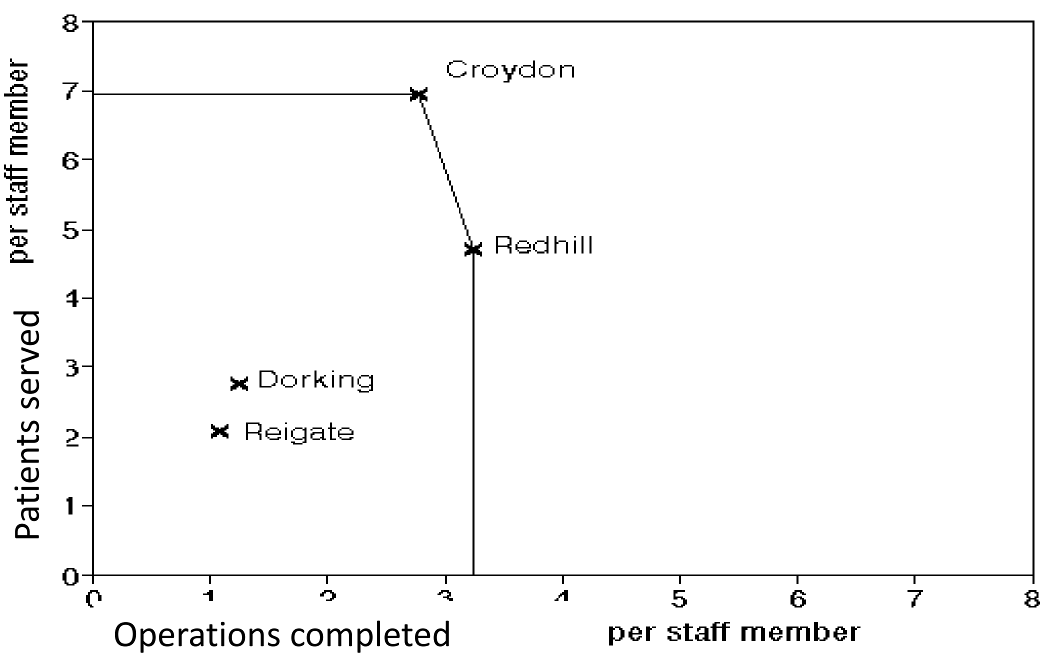
DEA only gives you *relative efficiencies* - efficiencies relative to the data considered. It does not, and cannot, give you absolute efficiencies.

We have used no new information here, merely taken data on inputs and outputs and presented it in a particular way.

Note too that the statement that a hospital has an efficiency of 100% is a strong statement, namely that we have no other hospital that can be said to be better than it.

Quantifying efficiency scores for inefficient DMU‘s

- Consider hospitals Dorking and Reigate
- With respect to both of the ratios Croydon (for example) **dominates** both Dorking and Reigate
- Plainly both Dorking and Reigate are less than 100% efficient.
- But how much? Can we assign an appropriate numerical value?



Consider Reigate, we have:

number of staff 110

patients served('000s) 230

patients served per staff member (230/110) = 2.09

operations completed('000s) 120

operations completed per staff member (120/110) = 1.09

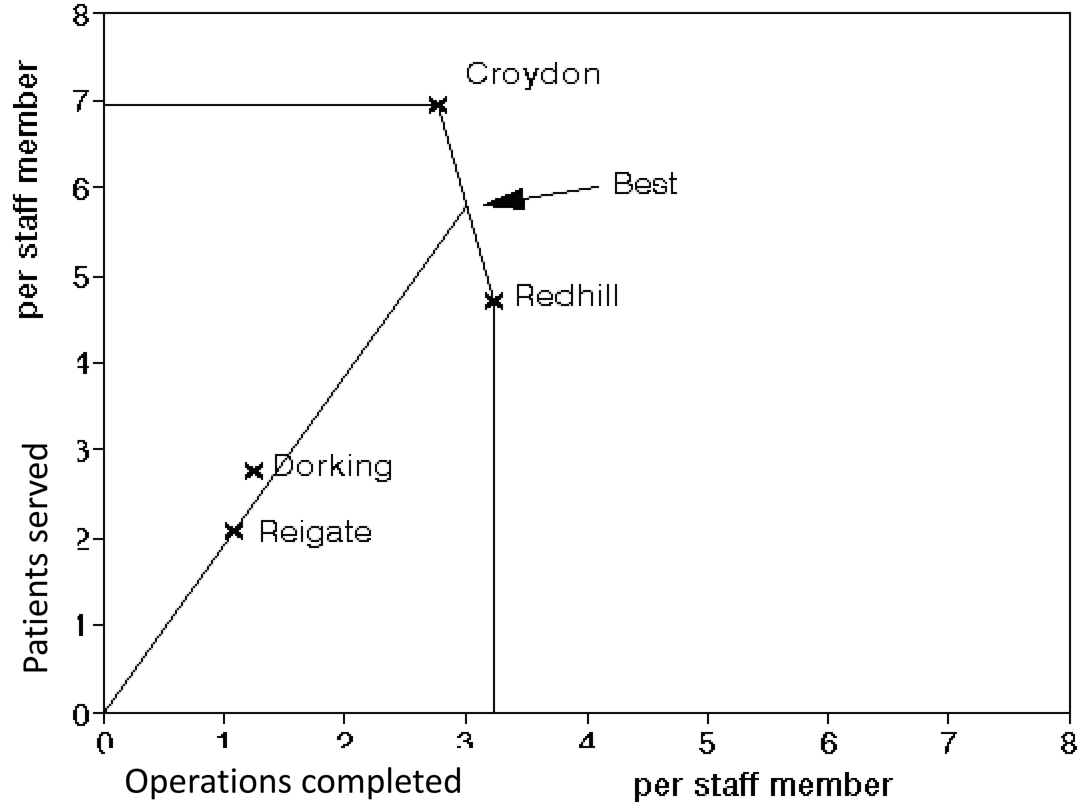
For Reigate the ratio patients served/operations completed

= (230/120) = 1.92

i.e. there are 1.92 patients served for operations completed.

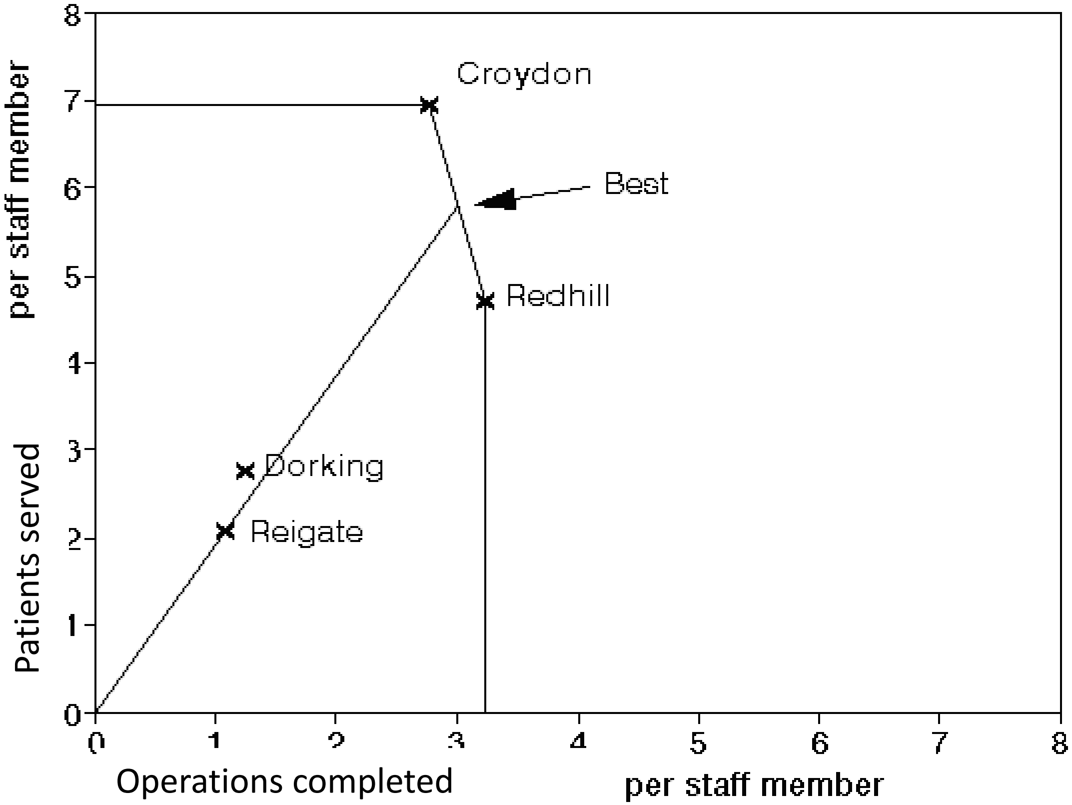
By definition this figure of 1.92 is also the ratio of:

- patients served per staff member: operations completed per staff member
- In other words (2.09/1.09) is also equal to 1.92



Any branch with a ratio (patients served per staff member: operations completed per staff member) equal to 1.92 lies on the straight line from the origin through Reigate
The slope (gradient) of this line is 1.92 - i.e. there are 1.92 patients served for every operation completed

Hence if Reigate were to retain the *same* business mix (i.e. 1.92 patients served for every operation completed) but to vary the number of staff it employs its performance would lie on the line from the origin through its current position as shown above.



It might seem reasonable to suggest therefore that the best possible performance that Reigate could be expected to achieve is given by the point labelled Best in the diagram above. This is the point where the line from the origin through Reigate meets the efficient frontier.

In other words Best represents a branch that, were it to exist, would have the same business mix as Reigate and would have an efficiency of 100%.

Then in DEA we numerically measure the (relative) efficiency of Reigate by the ratio:

100(length of line from origin to Reigate/length of line from origin through Reigate to efficient frontier)

For Reigate this is an efficiency of 36%.

The logic here is to compare the current performance of Reigate (the length of the line from the origin to Reigate) to the best possible performance that Reigate could reasonably be expected to achieve (the length of the line from the origin through Reigate to the efficient frontier).

Achieving the efficient frontier

The point labelled Best on the efficient frontier is considered to represent the best possible performance that Reigate can reasonably be expected to achieve.

There are a number of ways by which Reigate can move towards that point. It can:

•reduce its input (number of staff) whilst keeping its output constant (an input target)

•increase both its outputs, retaining the current ratio of 1.92 whilst keeping its input (number of staff) constant (an output target)

•do some combination of the above.

- In fact the same diagram we used to calculate the efficiency of Reigate can be used to set targets in a graphical manner.

- Suppose we say to Reigate that, next time period, their target is to achieve an efficiency of 40% (i.e. effectively a 10% increase in current efficiency).

- We know where, on the line from the origin to Best, a branch with the same business mix as Reigate but an efficiency of 40%, lies.

- We can say to Reigate that their goal is to move from their current position to that new position, and the combination of input/output changes necessary to achieve that is left up to them.

Use of the efficiencies

It is important to be clear about the appropriate use of the (relative) efficiencies we have calculated. Here we have:

Croydon 100%

Dorking 43%

Redhill 100%

Reigate 36%

This does NOT automatically mean that Reigate is only approximately one-third as efficient as the best branches.

Rather the efficiencies here would usually be taken as indicative of the fact that other branches are adopting practices and procedures which, if Reigate were to adopt them, would enable it to *improve* its performance.

This naturally invokes issues of highlighting and disseminating examples of best practice.

Equally there are issues relating to the identification of poor practice.

Use of the efficiencies

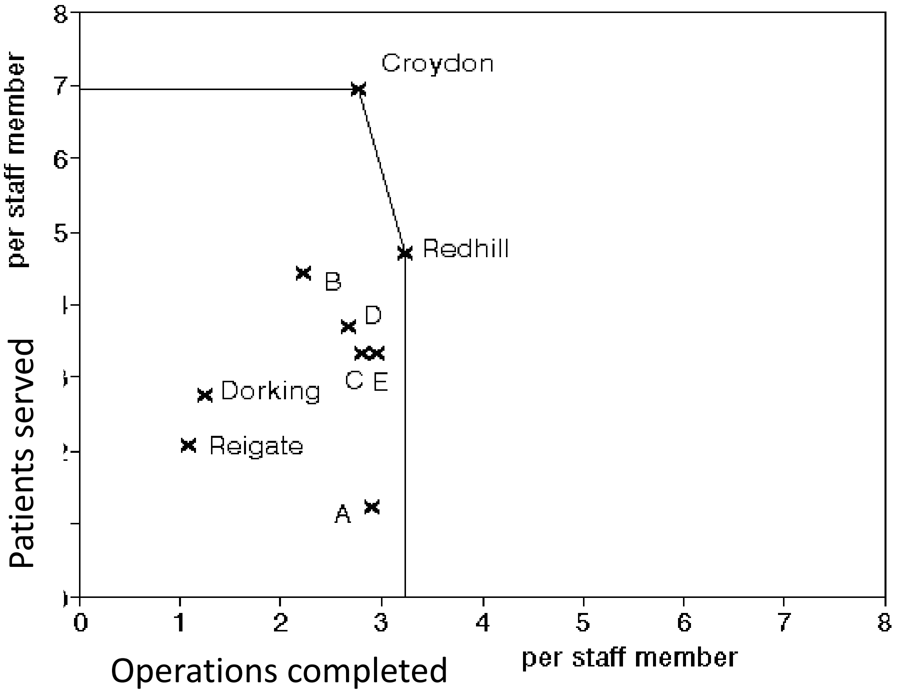
In DEA the concept of the *reference set* can be used to identify best performing branches with which to compare a poorly performing branch. Consider Reigate above.

The Best point associated with Reigate lies on the efficient frontier. A branch at this point would be the best possible branch to compare Reigate with as it would have the same business mix. No such branch exists however so we go to the two efficient branches either side of this Best point. These branches, Croydon and Redhill, are the **reference set** for Reigate.

The reference set for any branch with less than 100% efficiency consists of those branches with 100% efficiency with which it is most directly comparable.

Broadly put this means that the branches in the reference set have a "similar" mix of inputs and output.

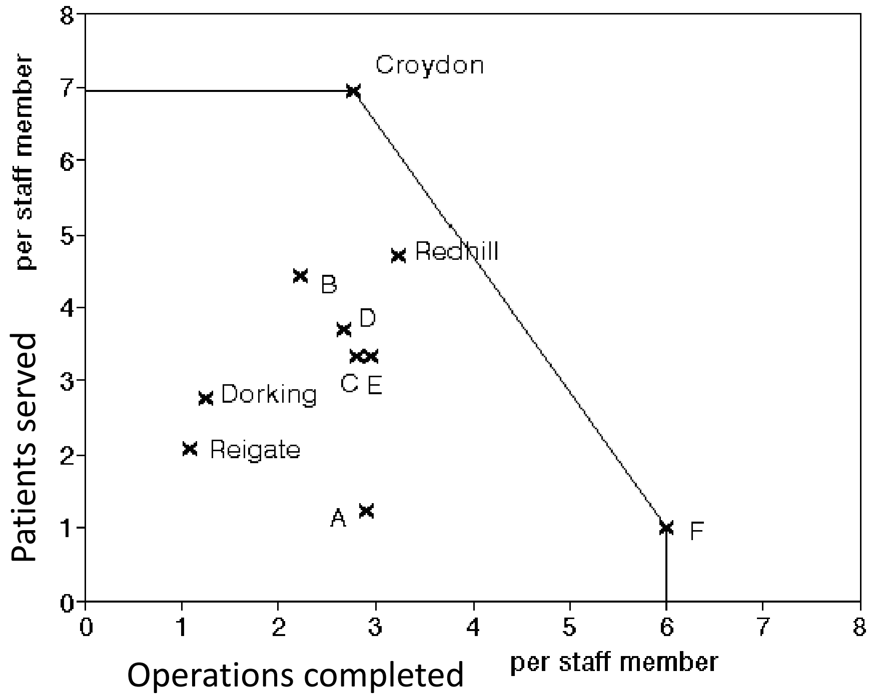
Including additional branches



What would be the reference sets for branches A to E?

Suppose that we add a branch F with patients served per staff member=1 and operations completed per staff member =6. What changes as a result of including this branch in the analysis?

Including additional branches



The efficient frontier now excludes Redhill.

We do not draw that efficient frontier from Croydon to Redhill and from Redhill to F for two reasons:

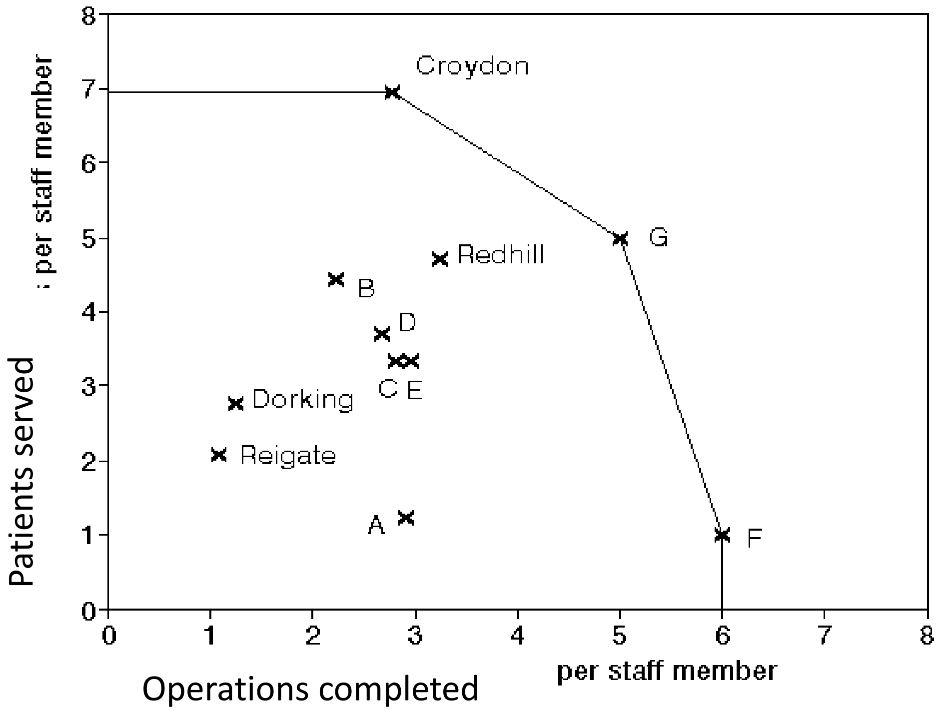
•mathematically the efficient frontier must be convex

•although we have not seen any branches on the line from Croydon to F it is assumed in DEA that we could construct *virtual branches*, which would be a linear combination of Croydon and F, and which would lie on the straight line from Croydon to F.

In the above it is clear why Croydon and F have a relative efficiency of 100% (i.e. are efficient), both are the top performers with respect to one of the two ratios we are considering.

A branch can be efficient even if it is not a top performer.

In the diagram below G is efficient since under DEA it is judged to have "strength with respect to both ratios", even though it is not the top performer in either.



Note here that in the above diagram the "feasible space" in which a single branch might lie and still achieve 100% efficiency without being a top performer in either ratio is quite large

We have shown how a simple graphical analysis of data on inputs and outputs can be used to calculate efficiencies.

Once such an analysis has been carried out then we can begin to tackle, with a clearer degree of insight than we had before, issues such as:

identification of best practice

identification of poor practice

target setting

resource allocation

monitoring efficiency changes over time

EXTENDING TO MORE INPUTS AND OUTPUTS

Extending to more inputs/outputs

- In our simple example we had just one input and two outputs. This is ideal for a simple graphical analysis.
- If we have more inputs or outputs than drawing simple pictures is not possible.
- However it is still possible to carry out EXACTLY the same analysis as above but using mathematics rather than pictures.

Extending to more inputs/outputs

<u>In evaluating any number of DMU's, with any number of inputs and outputs:</u>

- The inputs and outputs for each DMU should be specified;
- Define efficiency for each DMU as a weighted sum of outputs [total output] divided by a weighted sum of inputs [total input]
- All efficiencies are restricted to lie between zero and one (i.e. between 0% and 100%).
- In calculating the numerical value for the efficiency of a particular DMU weights are chosen so as maximise its efficiency, thereby presenting the DMU in the best possible light.

Discussion and model building in class