Chapter 12

Data Envelopment Analysis

Data Envelopment Analysis (DEA) is an increasingly popular management tool. This write-up is an introduction to Data Envelopment Analysis (DEA) for people unfamiliar with the technique. For a more in-depth discussion of DEA, the interested reader is referred to Seiford and Thrall [1990] or the seminal work by Charnes, Cooper, and Rhodes [1978].

DEA is commonly used to evaluate the efficiency of a number of producers. A typical statistical approach is characterized as a central tendency approach and it evaluates producers relative to an average producer In contrast, DEA compares each producer with only the "best" producers. By the way, in the DEA literature, a producer is usually referred to as a decision making unit or DMU. DEA is not always the right tool for a problem but is appropriate in certain cases. (See Strengths and Limitations of DEA.)

In DEA, there are a number of *producers*. The production process for each producer is to take a set of inputs and produce a set of outputs. Each producer has a varying level of inputs and gives a varying level of outputs. For instance, consider a set of banks. Each bank has a certain number of tellers, a certain square footage of space, and a certain number of managers (the inputs). There are a number of measures of the output of a bank, including number of checks cashed, number of loan applications processed, and so on (the outputs). DEA attempts to determine which of the banks are most efficient, and to point out specific inefficiencies of the other banks.

A fundamental assumption behind this method is that if a given producer, A, is capable of producing Y(A) units of output with X(A) inputs, then other producers should also be able to do the same if they were to operate efficiently. Similarly, if producer B is capable of producing Y(B) units of output with X(B) inputs, then other producers should also be capable of the same production schedule. Producers A, B, and others can then be combined to form a composite producer with composite inputs and composite outputs. Since this composite producer does not necessarily exist, it is typically called a virtual producer.

The heart of the analysis lies in finding the "best" virtual producer for each real producer. If the virtual producer is better than the original producer by either making more output with the same input or making the same output with less input then the original producer is inefficient. The subtleties of DEA are introduced in the various ways that producers A and B can be scaled up or down and combined.

12.1 Numerical Example

To illustrate how DEA works, let's take an example of three banks. Each bank has exactly 10 tellers (the only input), and we measure a bank based on two outputs: Checks cashed and Loan

applications. The data for these banks is as follows:

- Bank A: 10 tellers, 1000 checks, 20 loan applications
- Bank B: 10 tellers, 400 checks, 50 loan applications
- Bank C: 10 tellers, 200 checks, 150 loan applications

Now, the key to DEA is to determine whether we can create a virtual bank that is better than one or more of the real banks. Any such dominated bank will be an inefficient bank.

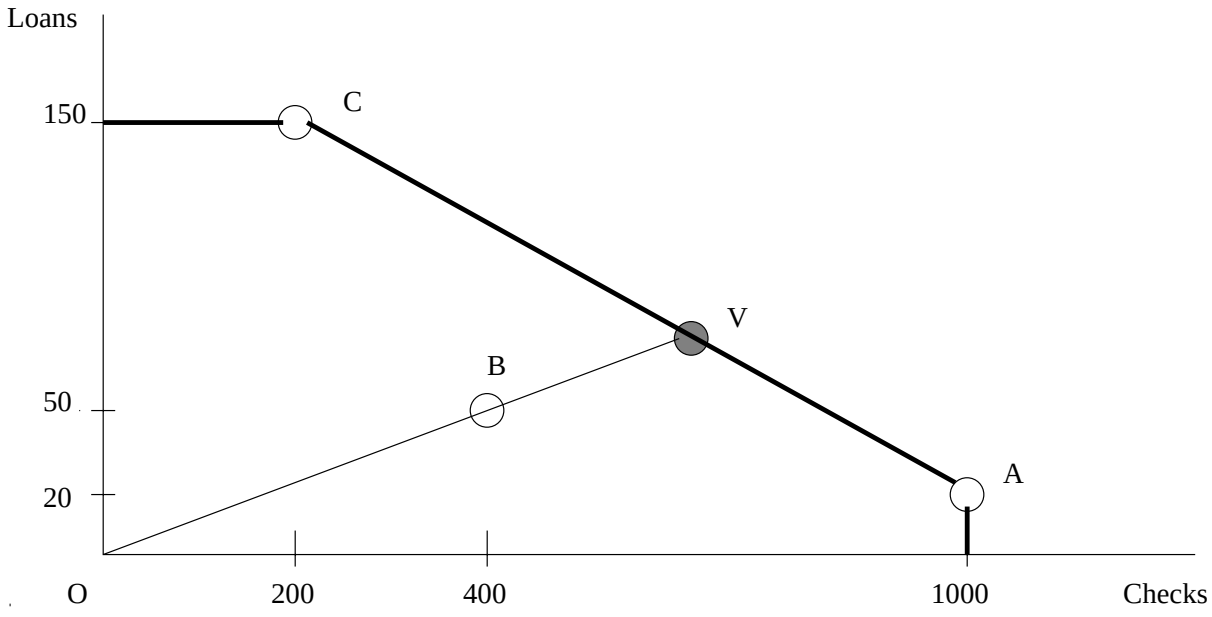
Consider trying to create a virtual bank that is better than Bank A. Such a bank would use no more inputs than A (10 tellers), and produce at least as much output (1000 checks and 20 loans). Clearly, no combination of banks B and C can possibly do that. Bank A is therefore deemed to be efficient. Bank C is in the same situation.

However, consider bank B. If we take half of Bank A and combine it with half of Bank C, then we create a bank that processes 600 checks and 85 loan applications with just 10 tellers. This dominates B (we would much rather have the virtual bank we created than bank B). Bank B is therefore inefficient.

Another way to see this is that we can scale down the inputs to B (the tellers) and still have at least as much output. If we assume (and we do), that inputs are linearly scalable, then we estimate that we can get by with 6.3 tellers. We do that by taking .34 times bank A plus .29 times bank B. The result uses 6.3 tellers and produces at least as much as bank B does. We say that bank B's efficiency rating is .63. Banks A and C have an efficiency rating of 1.

12.2 Graphical Example

The single input two-output or two input-one output problems are easy to analyze graphically. The previous numerical example is now solved graphically. (An assumption of constant returns to scale is made and explained in detail later.) The analysis of the efficiency for bank B looks like the following:



If it is assumed that convex combinations of banks are allowed, then the line segment connecting banks A and C shows the possibilities of virtual outputs that can be formed from these two banks. Similar segments can be drawn between A and B along with B and C. Since the segment AC lies beyond the segments AB and BC, this means that a convex combination of A and C will create the most outputs for a given set of inputs.

This line is called the efficiency frontier. The efficiency frontier defines the maximum combinations of outputs that can be produced for a given set of inputs.

Since bank B lies below the efficiency frontier, it is inefficient. Its efficiency can be determined by comparing it to a virtual bank formed from bank A and bank C. The virtual player, called V, is approximately 54% of bank A and 46% of bank C. (This can be determined by an application of the lever law. Pull out a ruler and measure the lengths of AV, CV, and AC. The percentage of bank C is then AV/AC and the percentage of bank A is CV/AC.)

The efficiency of bank B is then calculated by finding the fraction of inputs that bank V would need to produce as many outputs as bank B. This is easily calculated by looking at the line from the origin, O, to V. The efficiency of player B is OB/OV which is approximately 63%. This figure also shows that banks A and C are efficient since they lie on the efficiency frontier. In other words, any virtual bank formed for analyzing banks A and C will lie on banks A and C respectively. Therefore since the efficiency is calculated as the ratio of OA/OV or OA/OV, banks A and C will have efficiency scores equal to 1.0.

The graphical method is useful in this simple two dimensional example but gets much harder in higher dimensions. The normal method of evaluating the efficiency of bank B is by using an linear programming formulation of DEA.

Since this problem uses a constant input value of 10 for all of the banks, it avoids the complications caused by allowing different returns to scale. Returns to scale refers to increasing or decreasing efficiency based on size. For example, a manufacturer can achieve certain economies of scale by producing a thousand circuit boards at a time rather than one at a time - it might be only 100 times as hard as producing one at a time. This is an example of increasing returns to scale (IRS.)

On the other hand, the manufacturer might find it more than a trillion times as difficult to produce a trillion circuit boards at a time though because of storage problems and limits on the worldwide copper supply. This range of production illustrates decreasing returns to scale (DRS.) Combining the two extreme ranges would necessitate variable returns to scale (VRS.)

Constant Returns to Scale (CRS) means that the producers are able to linearly scale the inputs and outputs without increasing or decreasing efficiency. This is a significant assumption. The assumption of CRS may be valid over limited ranges but its use must be justified. As an aside, CRS tends to lower the efficiency scores while VRS tends to raise efficiency scores.

12.3 Using Linear Programming

Data Envelopment Analysis, is a linear programming procedure for a frontier analysis of inputs and outputs. DEA assigns a score of 1 to a unit only when comparisons with other relevant units do not provide evidence of inefficiency in the use of any input or output. DEA assigns an efficiency score less than one to (relatively) inefficient units. A score less than one means that a linear combination of other units from the sample could produce the same vector of outputs using a smaller vector of inputs. The score reflects the radial distance from the estimated production frontier to the DMU under consideration.

There are a number of equivalent formulations for DEA. The most direct formulation of the exposition I gave above is as follows:

Let X_i be the vector of inputs into DMU i. Let Y_i be the corresponding vector of outputs. Let X_0 be the inputs into a DMU for which we want to determine its efficiency and Y_0 be the outputs. So the X's and the Y's are the data. The measure of efficiency for DMU_0 is given by the following linear program:

$$\begin{array}{lrcl} Min & \theta & & \\ & & & \\ s.t. & \sum \lambda_i X_i & \leq & \theta X_0 \\ & \sum \lambda_i Y_i & \geq & Y_0 \\ & \lambda \geq 0 & & \end{array}$$

where λ_i is the weight given to DMU i in its efforts to dominate DMU 0 and θ is the efficiency of DMU 0. So the λ's and θ are the variables. Since DMU 0 appears on the left hand side of the equations as well, the optimal θ cannot possibly be more than 1. When we solve this linear program, we get a number of things:

1. The efficiency of DMU 0 (θ), with $\theta = 1$ meaning that the unit is efficient.
2. The unit's "comparables" (those DMU with nonzero λ).
3. The "goal" inputs (the difference between X_0 and $\sum \lambda_i X_i$)
4. Alternatively, we can keep inputs fixed and get goal outputs ($\frac{1}{\theta} \sum_i Y_i$)

DEA assumes that the inputs and outputs have been correctly identified. Usually, as the number of inputs and outputs increase, more DMUs tend to get an efficiency rating of 1 as they become too specialized to be evaluated with respect to other units. On the other hand, if there are too few inputs and outputs, more DMUs tend to be comparable. In any study, it is important to focus on correctly specifying inputs and outputs.

Example 12.3.1 *Consider analyzing the efficiencies of 3 DMUs where 2 inputs and 3 outputs are used. The data is as follows:*

DMU	*Inputs*		*Outputs*		
1	*5*	*14*	*9*	*4*	*16*
2	*8*	*15*	*5*	*7*	*10*
3	*7*	*12*	*4*	*9*	*13*

The linear programs for evaluating the 3 DMUs are given by:

- **LP for evaluating DMU 1:**

```
min THETA
st
5L1+8L2+7L3 - 5THETA <= 0
14L1+15L2+12L3 - 14THETA <= 0
9L1+5L2+4L3 >= 9
4L1+7L2+9L3 >= 4
16L1+10L2+13L3 >= 16
L1, L2, L3 >= 0
```

- **LP for evaluating DMU 2:**

```
min THETA
st
5L1+8L2+7L3 - 8THETA <= 0
14L1+15L2+12L3 - 15THETA <= 0
9L1+5L2+4L3 >= 5
4L1+7L2+9L3 >= 7
16L1+10L2+13L3 >= 10
L1, L2, L3 >= 0
```

- **LP for evaluating DMU 3:**

```
min THETA
st
5L1+8L2+7L3 - 7THETA <= 0
14L1+15L2+12L3 - 12THETA <= 0
9L1+5L2+4L3 >= 4
4L1+7L2+9L3 >= 9
16L1+10L2+13L3 >= 13
L1, L2, L3 >= 0
```

The solution to each of these is as follows:

- **DMU 1.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
B10	theta	1	0	1	1E+30	1
B11	L1	1	0	0	0.92857142	0.619047619
B12	L2	0	0.24285714	0	1E+30	0.242857143
B13	L3	0	0	0	0.36710963	0.412698413

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
B16	IN1	-0.103473	0	0	1E+30	0
B17	IN2	-0.289724	-0.07142857	0	0	1E+30
B18	OUT1	9	0.085714286	9	0	0
B19	OUT2	4	0.057142857	4	0	0
B20	OUT3	16	0	16	0	1E+30

- **DMU 2.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
B10	theta	0.773333333	0	1	1E+30	1
B11	L1	0.261538462	0	0	0.866666667	0.577777778
B12	L2	0	0.22666667	0	1E+30	0.226666667
B13	L3	0.661538462	0	0	0.342635659	0.385185185

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
B16	IN1	-0.24820512	0	0	1E+30	0.248205128
B17	IN2	-0.452651	-0.0666667	0	0.46538461	1E+30
B18	OUT1	5	0.08	5	10.75	0.655826558
B19	OUT2	7	0.0533333	7	1.05676855	3.41509434
B20	OUT3	12.78461538	0	10	2.78461538	1E+30

- **DMU 3.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
B10	theta	1	0	1	1E+30	1
B11	L1	0	0	0	1.08333333	0.722222222
B12	L2	0	0.283333333	0	1E+30	0.283333333
B13	L3	1	0	0	0.42829457	0.481481481

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
B16	IN1	-0.559375	0	0	1E+30	0
B17	IN2	-0.741096	-0.08333333	0	0	1E+30
B18	OUT1	4	0.1	4	16.25	0
B19	OUT2	9	0.066666667	9	0	0
B20	OUT3	13	0	13	0	1E+30

Note that DMUs 1 and 3 are overall efficient and DMU 2 is inefficient with an efficiency rating of 0.773333.

Hence the efficient levels of inputs and outputs for DMU 2 are given by:

- Efficient levels of Inputs:

$$0.261538 \begin{bmatrix} 5 \\ 14 \end{bmatrix} + 0.661538 \begin{bmatrix} 7 \\ 12 \end{bmatrix} = \begin{bmatrix} 5.938 \\ 11.6 \end{bmatrix}$$

- Efficient levels of Outputs:

$$0.261538 \begin{bmatrix} 9 \\ 4 \\ 16 \end{bmatrix} + 0.661538 \begin{bmatrix} 4 \\ 9 \\ 13 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 12.785 \end{bmatrix}$$

Note that the outputs are at least as much as the outputs currently produced by DMU 2 and inputs are at most as big as the 0.773333 times the inputs of DMU 2. This can be used in two different ways: The inefficient DMU should target to cut down inputs to equal at most the efficient levels. Alternatively, an equivalent statement can be made by finding a set of efficient levels of inputs and outputs by dividing the levels obtained by the efficiency of DMU 2. This focus can then be used to set targets primarily for outputs rather than reduction of inputs.

Alternate Formulation

There is another, probably more common formulation, that provides the same information. We can think of DEA as providing a price on each of the inputs and a value for each of the outputs. The efficiency of a DMU is simply the ratio of the inputs to the outputs, and is constrained to be no more than 1. The prices and values have nothing to do with real prices and values: they are an artificial construct. The goal is to find a set of prices and values that puts the target DMU in the best possible light. The goal, then is to

$$\begin{array}{lll} Max & \frac{u^T Y_0}{v^T X_0} & \\ & & \\ s.t. & \frac{u^T Y_j}{v^T X_j} \leq 1, & j = 0, ..., n, \\ & & \\ & u^T \geq 0, & \\ & v^T \geq 0. & \end{array}$$

Here, the variables are the u's and the v's. They are vectors of prices and values respectively.

This fractional program can be equivalently stated as the following linear programming problem (where Y and X are matrices with columns Y_j and X_j respectively).

$$\begin{array}{lll} Max & u^T Y_0 & \\ & & \\ s.t. & v^T X_0 = 1, & \\ & & \\ & u^T Y - v^T X \leq 0, & \\ & & \\ & u^T \geq 0, & \\ & v^T \geq 0. & \end{array}$$

We denote this linear program by (D). Let us compare it with the one introduced earlier, which we denote by (P):

$$\begin{array}{lll} Min & \theta & \\ & & \\ s.t. & \sum \lambda_i X_i \leq \theta X_0 & \\ & \sum \lambda_i Y_i \geq Y_0 & \\ & \lambda \geq 0. & \end{array}$$

To fix ideas, let us write out explicitely these two formulations for DMU 2, say, in our example.

Formulation (P) for DMU 2:

```
min                      THETA
st
-5 L1 - 8 L2 - 7 L3 + 8 THETA >= 0
-14L1 -15 L2 -12 L3 + 15THETA >= 0
 9 L1 + 5 L2 + 4 L3           >= 5
 4 L1 + 7 L2 + 9 L3           >= 7
 16L1 +10 L2 +13 L3           >= 10
L1>=0, L2>=0, L3>=0
```

Formulation (D) for DMU 2:

```
max             5 U1 + 7 U2 + 10 U3
st
- 5 V1 - 14V2 + 9 U1 + 4 U2 + 16 U3 <= 0
- 8 V1 - 15V2 + 5 U1 + 7 U2 + 12 U3 <= 0
- 7 V1 - 12V2 + 4 U1 + 9 U2 + 13 U3 <= 0
  8 V1 + 15V2                        = 1
    V1>=0, V2>=0, U1>=0, U2>=0,  U3>=0
```

Formulations (P) and (D) are dual linear programs! These two formulations actually give the same information. You can read the solution to one from the shadow prices of the other. We will not discuss linear programming duality in this course. You can learn about it in some of the OR electives.

Exercise 85 Consider the following baseball players:

Name	At Bat	Hits	HomeRuns
Martin	135	41	6
Polcovich	82	25	1
Johnson	187	40	4

(You need know nothing about baseball for this question). In order to determine the efficiency of each of the players, At Bats is defined as an input while Hits and Home Runs are outputs. Consider the following linear program and its solution:

```
MIN     THETA
SUBJECT TO
          - 187 THETA + 135 L1 + 82 L2 + 187 L3 <=   0
           41 L1 + 25 L2 + 40 L3 >=   40
            6 L1 + L2 + 4 L3 >=   4
              L1, L2, L3 >= 0
```

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient
B10	THETA	0.703135	0	1
B11	L1	0.550459	0	0
B12	L2	0.697248	0	0
B13	L3	0	0.296865	0

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side
B16	AT BATS	0	-0.005348	0
B17	HITS	0	0.017515	40
B18	HOME RUNS	0	0.000638	4

(a) For which player is this a DEA analysis? Is this player efficient? What is the efficiency rating of this player? Give the "virtual producer" that proves that efficiency rating (you should give the At bats, Hits, and Home Runs for this virtual producer).

(b) Formulate the linear program for Jonhson using the alternate formulation and solve using Solver. Compare the "Final Value" and "Shadow Price" columns from your Solver output with the solution given above.

12.4 Applications

The simple bank example described earlier may not convey the full view on the usefulness of DEA. It is most useful when a comparison is sought against "best practices" where the analyst doesn't want the frequency of poorly run operations to affect the analysis. DEA has been applied in many situations such as: health care (hospitals, doctors), education (schools, universities), banks, manufacturing, benchmarking, management evaluation, fast food restaurants, and retail stores.

The analyzed data sets vary in size. Some analysts work on problems with as few as 15 or 20 DMUs while others are tackling problems with over 10,000 DMUs.

12.5 Strengths and Limitations of DEA

As the earlier list of applications suggests, DEA can be a powerful tool when used wisely. A few of the characteristics that make it powerful are:

- DEA can handle multiple input and multiple output models.
- It doesn't require an assumption of a functional form relating inputs to outputs.
- DMUs are directly compared against a peer or combination of peers.
- Inputs and outputs can have very different units. For example, X1 could be in units of lives saved and X2 could be in units of dollars without requiring an a priori tradeoff between the two.

The same characteristics that make DEA a powerful tool can also create problems. An analyst should keep these limitations in mind when choosing whether or not to use DEA.

- Since DEA is an extreme point technique, noise (even symmetrical noise with zero mean) such as measurement error can cause significant problems.
- DEA is good at estimating "relative" efficiency of a DMU but it converges very slowly to "absolute" efficiency. In other words, it can tell you how well you are doing compared to your peers but not compared to a "theoretical maximum."
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research.
- Since a standard formulation of DEA creates a separate linear program for each DMU, large problems can be computationally intensive.

12.6 References

DEA has become a popular subject since it was first described in 1978. There have been hundreds of papers and technical reports published along with a few books. Technical articles about DEA have been published in a wide variety of places making it hard to find a good starting point. Here are a few suggestions as to starting points in the literature.

1. Charnes, A., W.W. Cooper, and E. Rhodes. "Measuring the efficiency of decision making units." European Journal of Operations Research (1978): 429-44.
2. Banker, R.D., A. Charnes, and W.W. Cooper. "Some models for estimating technical and scale inefficiencies in data envelopment analysis." Management Science 30 (1984): 1078-92.
3. Dyson, R.G. and E. Thanassoulis. "Reducing weight flexibility in data envelopment analysis." Journal of the Operational Research Society 39 (1988): 563-76.
4. Seiford, L.M. and R.M. Thrall. "Recent developments in DEA: the mathematical programming approach to frontier analysis." Journal of Econometrics 46 (1990): 7-38.
5. Ali, A.I., W.D. Cook, and L.M. Seiford. "Strict vs. weak ordinal relations for multipliers in data envelopment analysis." Management Science 37 (1991): 733-8.
6. Andersen, P. and N.C. Petersen. "A procedure for ranking efficient units in data envelopment analysis." Management Science 39 (1993): 1261-4.
7. Banker, R.D. "Maximum likelihood, consistency and data envelopment analysis: a statistical foundation." Management Science 39 (1993): 1265-73.

The first paper was the original paper describing DEA and results in the abbreviation CCR for the basic constant returns-to-scale model. The Seiford and Thrall paper is a good overview of the literature. The other papers all introduce important new concepts. This list of references is certainly incomplete.

A good source covering the field of productivity analysis is The Measurement of Productive Efficiency edited by Fried, Lovell, and Schmidt, 1993, from Oxford University Press. There is also

a recent book from Kluwer Publishers, Data Envelopment Analysis: Theory, Methodology, and Applications by Charnes, Cooper, Lewin, and Seiford.

To stay more current on the topics, some of the most important DEA articles appear in Management Science, The Journal of Productivity Analysis, The Journal of the Operational Research Society, and The European Journal of Operational Research. The latter just published a special issue, "Productivity Analysis: Parametric and Non-Parametric Approaches" edited by Lewin and Lovell which has several important papers.