Data Envelopment Analysis and its aplications

Lucie Chytilová

VŠB-Technical University of Ostrava
Czech Republic

13. Letní škola aplikované informatiky Bedřichov

DEA

Content

Introduction

The aim

to calculate and analyze the efficiency of homogenous DMUs

- Banks in the V4 countries
- Colleges in Great Britain
- Erasmus students in VŠB - TU Ostrava
- Metallurgy - emission trading
- Cooperation in Project Management
- IT Projects in Poland

Overview of the Literature

- Charnes, Cooper and Rhodes (1978)
  - Measuring the efficiency of decision making units
- Banker, Charnes and Cooper (1984)
  - Extension of basic models

Data Envelopment Analysis

- Non-parametric technique based on linear programming for measuring the relative efficiency of a set of DMUs
- DEA models differs according to orientation or assumptions on return to scale
- An essential topic is how to choose the appropriate inputs and outputs
- The analysis of non-efficient units, projection of non-efficient units on the frontier etc.
- There exist many improvements of the classical DEA models - quasi fixed variables, fuzzy variables etc.

DEA

- the non-parametric approach for measuring the relative efficiency of the number of DMUs,
- let DMU_k for $k = 1, 2, ..., T$ and let input and output data for DMU_k be $\mathbf{X} = \{x_{ik}, i = 1, 2, ...R; k = 1, 2, ...T\}$ and $\mathbf{Y} = \{y_{jk}, j = 1, 2, ...S; k = 1, 2, ...T\}$, u_i for $i = 1, 2, ...R$ and v_j for $j = 1, 2, ...S$ be the weights of i^{th} input and j^{th} output, respectively,
- the relative efficiency score of DMU_k can be define as:

$$e_k = \frac{\sum_{j=1}^{S} v_j y_{jk}}{\sum_{i=1}^{R} u_i x_{ik}}, \text{ for } k = 1, 2, ..., T. \qquad (1)$$

The multiplier input-oriented model - CCR

- Charmes et al. (1978),
- to measure the efficiency score of the under evaluation unit, DMU_Q where $Q \in \{1, ..., T\}$:

$$
\begin{array}{lll}
 & \max\ e_Q = \sum_{j=1}^{S} v_j y_{jQ}, & \\
\text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\
 & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, 2, ..., T, \\
 & u_i \geq 0, & i = 1, 2, ..., R, \\
 & v_j \geq 0, & j = 1, 2, ..., S.
\end{array}
\tag{2}
$$

- DMU_Q is CCR-efficient if and only if $e^* = 1$ and if there exists at least one optimal solution $(\overrightarrow{u^*}, \overrightarrow{v^*})$ with $\overrightarrow{u^*} > \overrightarrow{0}$ and $\overrightarrow{v^*} > \overrightarrow{0}$,
- inefficient units have a degree of relative efficiency less than one.

The multiplier input-oriented model - BCC

- Banker et al. (1984) extended the CCR model
- Convex envelope of data which leads to more efficient DMUs
- BCC model in dual multiplier form is mathematically as it follows:

$$\begin{array}{lll} & \max\ e_Q = \sum_{j=1}^{S} v_j y_{jQ} - v_0, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} - v_0 \leq 0, & k = 1, ..., T, \\ & u_i \geq 0, & i = 1, ..., R, \\ & v_j \geq 0, & j = 1, ..., S, \\ & v_0 \in (-\infty, \infty), & \end{array} \tag{3}$$

where v_0 is the dual variable assigned to the convexity condition $\mathbf{e^T}\lambda = \mathbf{1}$ of envelopment form of BCC model

Imprecise data envelopment analysis - IDEA

- if x_{ik}, y_{jk} are imprecise and unknown decision variables as bounded and ordinal data, Cook et al. (1993, 1996), then model (2) becomes a non-linear and non-convex program;
- it is called imprecise DEA (IDEA), Zhu (2002, 2003);

Imprecise data envelopment analysis - IDEA

- weak ordinal data can be expressed as

$$y_{jk} \leq y_{jl} \text{ and } x_{ik} \leq x_{il} \quad \forall k \neq l \tag{4}$$

for $j \in DO$ and $i \in DI$, or to simplify the presentation

$$y_{j1} \leq y_{j2} \leq \ldots \leq y_{jl} \leq \ldots \leq y_{jn} \quad (j \in DO), \tag{5}$$
$$x_{i1} \leq x_{i2} \leq \ldots \leq x_{il} \leq \ldots \leq x_{in} \quad (i \in DI),$$

where DO and DI represents the associated sets containing weak ordinal outputs and inputs, respectively.

- the strong ordinal data can be expressed as:

$$y_{j1} < y_{j2} < \ldots < y_{jl} < \ldots < y_{jn} \quad (j \in SO), \tag{6}$$
$$x_{i1} < x_{i2} < \ldots < x_{il} < \ldots < x_{in} \quad (i \in SI),$$

where SO and SI represents the associated sets containing weak ordinal outputs and inputs, respectively.

Imprecise data envelopment analysis - IDEA

- model (2) involving (3)-(5) then changes into the following model:

$$\begin{array}{lll} & \max\ e_Q = \sum_{j=1}^{S} v_j y_{jQ}, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, 2, ..., T, \\ & (x_{ik}) \in \theta_i^-, & i = 1, 2, ..., R, \\ & (y_{jk}) \in \theta_j^+, & j = 1, 2, ..., S, \\ & u_i \geq 0, v_j \geq 0, & \end{array} \qquad (7)$$

  where $x_{ik} \in \theta_i^-$ and $y_{jk} \in \theta_j^+$ represents any of or all of (3)-(5);
- can be solved by the standard linear DEA models vis concerning the bounded and ordinal data into exact data, Zhu (2002);

Imprecise data envelopment analysis - IDEA

- model (6) when θ_i^- and θ_j^+ are in forms of (4) and obtain a set of optimal solutions y_{jk}^* and x_{jk}^* with the optimal e_Q^* is the following:

$$
\begin{array}{ll}
 & e_Q^* = \max \sum_{j \in DO} v_j y_{jQ} + \sum_{j \notin DO} v_j y_{jQ}, \\
\text{s.t.} & \sum_{i \in DI} u_i x_{iQ} + \sum_{i \notin DI} u_i x_{iQ} = 1, \\
 & \sum_{j \in DO} v_j y_{jk} + \sum_{j \notin DO} v_j y_{jk} - \sum_{i \in DI} u_i x_{ik} \\
 & -\sum_{i \notin DI} u_i x_{ik} \leq 0, \\
 & k = 1, 2, ..., T, \\
 & u_i \geq 0, v_j \geq 0, \qquad i = 1, 2, ..., R, j = 1, 2, ..., S, \\
\text{where} & 0 \leq y_{j1}^* \leq y_{j2}^* \leq ... \leq y_{jl}^* \leq ... \leq y_{jn}^* \leq M \quad (j \in DO), \\
 & 0 \leq x_{i1}^* \leq x_{i2}^* \leq ... \leq x_{il}^* \leq ... \leq x_{in}^* \leq M \quad (i \in DI).
\end{array}
\tag{8}
$$

  where M is sufficiently large;

- linear CCR model;

Input and Output Variables

- Identification of relevant variables - classic DEA
  - Inputs - min
  - Outputs - max

DEA	Criteria	Max	Min	Averag
Input				
	overrun on budget	2.500	0.833	1.238
	overrun on schedule	0.500	2.667	1.435
Output				
	percentage of scope realized	1.000	0.850	0.961
	customer satsifaction	2.000	4.000	3.091

Table 1: Input and output variables with description

Input and Output Variables

- Identification of relevant variables - balanced scorecard (BSC)
  - Strategic planning and management system:
    - Financial perspective
    - Customer perspective
    - Internal-business-process perspective
    - Learning and growth perspective

DEA	BSC perspectives	Criteria
Output	Financial perspective	planned budget
Output		actual budget
Input	Customer perspective	planned price
Input		actual price
Input	Internal-business-process perspective	risk documentation
Input		type of document
Output	Learning and growth perspective	experience of manager
Input		number of people

Table 2: Input and output variables with description

Results and Discussions

#IT	classic DEA	I-DEA	I-DEA-BSC	#IT	classic DEA	I-DEA	I-DEA-BSC
1	0.470	0.525	0.618	37	0.390	0.429	0.474
2	0.366	0.413	0.443	38	0.807	0.934	1.000
3	0.683	0.870	0.969	39	0.986	1.000	1.000
4	0.640	0.697	0.776	40	0.910	0.923	1.000
5	0.566	0.640	0.686	41	0.805	0.933	1.000
6	0.448	0.493	0.533	42	0.780	0.870	0.969
7	0.725	0.847	0.917	43	0.921	1.000	1.000
8	0.657	0.762	0.818	44	0.823	0.875	0.976
9	0.429	0.756	0.848	45	0.659	0.732	0.789
10	0.706	0.824	0.980	46	0.869	0.910	1.000
11	0.623	0.712	0.761	47	0.513	0.691	0.778
12	0.780	0.917	1.000	48	0.780	0.917	1.000

Figure 1: Results of DEA

Results and Discussions

	classic DEA	I-DEA	I-DEA-BSC
max	1	1	1
min	0.37	0.41	0.44
std. dev	0.14	0.14	0.15
average	0.69	0.79	0.86

Table 3: Descriptive statistics for the results

Conclusion

- I-DEA-BSC is the best one
- Future work
  - Sets of inputs and outputs should be extended
  - DMUs should be revised, extend or divided into groups based on certain rules
  - Different version of the model may be established

Thank you for your attention!