

Data Envelopment Analysis with Partially Perfect Objects

Alexander Y. Vaninsky

Abstract—This paper presents a simplified version of Data Envelopment Analysis (DEA) - a conventional approach to evaluating the performance and ranking of competitive objects characterized by two groups of factors acting in opposite directions: inputs and outputs. DEA with a Perfect Object (DEA PO) augments the group of actual objects with a virtual Perfect Object - the one having greatest outputs and smallest inputs. It allows for obtaining an explicit analytical solution and making a step to an absolute efficiency. This paper develops this approach further and introduces a DEA model with Partially Perfect Objects. DEA PPO consecutively eliminates the smallest relative inputs or greatest relative outputs, and applies DEA PO to the reduced collections of indicators. The partial efficiency scores are combined to get the weighted efficiency score. The computational scheme remains simple, like that of DEA PO, but the advantage of the DEA PPO is taking into account all of the inputs and outputs for each actual object. Firm evaluation is considered as an example.

Keywords—Data Envelopment Analysis, Perfect object, Partially perfect object, Partial efficiency, Explicit solution, Simplified algorithm.

1. INTRODUCTION

DATA Envelopment Analysis (DEA) was developed in publications [1], [2]; its comprehensive description may be found in [3] and on the website <http://deazone.com/>. DEA estimates relative efficiencies of objects in a group, referred to as Decision - Making Units (DMUs) that use inputs $X = (X_j, j = 1, \dots, r)$ to produce outputs $Y = (Y_i, i = 1, \dots, s)$.

DEA combines all indicators into a single efficiency score scaled between 0 and 1. Efficient objects receive the score equal 1, inefficient objects, less than 1. To measure the efficiency, DEA uses the efficiency ratio suggested in [4]:

$$E = \frac{\sum_{k=1}^s u_k Y_k}{\sum_{l=1}^r v_l X_l} \quad (1)$$

where $u = (u_1, \dots, u_s)$ and $v = (v_1, \dots, v_r)$ are non-negative weights assigned to outputs and inputs, respectively.

DEA is a non-parametric method; it does not require a prior functional relationship between inputs and outputs and

the efficiency score. The main advantage of DEA is its ability to assign values to u and v objectively by solving a series of linear programming problems. To calculate an efficiency score, DEA allows each DMU to assign its own weight coefficients to each input and output favorably. However, the ability of a given DMU to achieve maximal possible efficiency score is restricted by the requirement that with the weight coefficients assigned by any given DMU to itself, no one other DMU in the group received an efficiency score greater than one. This means that a poorly performing DMU cannot achieve a high efficiency score for itself by playing with the weight coefficients, since an object that performs really well would have received the efficiency score greater than one.

The basic efficiency ratio (1) therefore, generates the following series of optimization problems:

For each DMU_{*i*}, $i = 1, \dots, n$, find non-negative vectors $u_i = (u_{i1}, \dots, u_{is})$ and $v_i = (v_{i1}, \dots, v_{ir})$ such that:

$$\text{maximize } E_i = \frac{\sum_{k=1}^s u_{ik} Y_{ik}}{\sum_{p=1}^r v_{ip} X_{ip}} \quad (2)$$

subject to

$$E_j \leq 1 \text{ with all } u_i = (u_{i1}, \dots, u_{is}), v_i = (v_{i1}, \dots, v_{ir}), i, j = 1, \dots, n. \quad (3)$$

The DEA model given by (2) and (3) is referred to as the ratio DEA model.

In the envelope DEA model, the set of optimization problems is changed for a set of equivalent linear programming ones:

For each DMU_{*i*}, $i = 1, \dots, n$, find a non-negative vector $\lambda_i = (\lambda_{i1}, \lambda_{i2}, \dots, \lambda_{in})$ and a scalar ω_i such that

$$\text{minimize } \omega_i$$

subject to

$$\begin{aligned} \sum_{j=1}^n \lambda_{ij} X_{jk} &\leq \omega_i X_{ik}, k = 1, \dots, r; \\ \sum_{j=1}^n \lambda_{ij} Y_{jp} &\geq Y_{ip}, p = 1, \dots, s; \\ \lambda_{ij} &\geq 0, j = 1, \dots, n; \\ 0 &< \omega_i \leq 1 \end{aligned} \quad (4)$$

A. Y. Vaninsky is with the Hostos Community College of The City University of New York, Bronx, NY 10451, USA (phone: 1-718-319-7930; fax: 1-718-518-6706; e-mail: avaninsky@hostos.cuny.edu).

where X_{jk} and Y_{jp} stand for the k -th input and p -th output of a DMU_j, respectively.

The LP-problem stated by (4) has the following interpretation: for each DMU_i, DEA-algorithm designs a virtual object that produces at least the same outputs as DMU_i while using at most ω_i - share of its inputs. This virtual DMU is constructed of λ_i - multiples of all DMUs, including the DMU_i itself. This LP-problem has at least one feasible solution:

$$\omega_i = 1, \lambda_{ii} = 1, \lambda_{ij} = 0 \text{ for } i \neq j \quad (5)$$

which means that a virtual DMU is the same as the DMU_i itself. For some DMUs this is the only solution, meaning that their performance cannot be improved by simulating peer DMUs. For other DMUs in the group, better solutions exist with a smaller value of $\omega_i < 1$. Such DMUs may perform better by acquiring the properties of their peers. The minimal value of ω_i given by LP-problem (4) and efficiency score E_i corresponding to the problem (2), (3) are equal:

$$\max E_i = \min \omega_i \quad (6)$$

DMUs with $E_i = 1$ are called efficient, otherwise, inefficient. An input-minimization DEA with constant returns to scale (IM CRS) is a natural extension of an intuitively clear formula (1) and possesses some useful properties. First, the efficiency scores remain the same if the input-minimization model is changed for the output - maximization one (OM CRS). Thus, a choice of a basic model becomes unambiguous. Second, the efficiency scores preserve their values if one or several inputs or outputs are changed proportionally. This is an important issue because some indicators may have units of measurement.

The conventional DEA has, however, at least three weaknesses. First, it measures a relative efficiency only. An object that has a high efficiency score in one group may or may not be equally efficient in another group. Second, the DEA may assign the fully efficient score to an object that has just one very large output or very small input. Third, the Linear Programming algorithm that DEA uses, is actually a black box with regard to the weight coefficients in the ratio DEA model (2). Though, theoretically, DEA takes into consideration all of the indicators by maximizing the ratio of the weighted sum of all outputs to the weighted sum of all inputs, the DEA optimal solution typically assigns the non-zero weights only to some of them. This makes it possible to increase inputs or decrease outputs having zero weights arbitrarily, without any change in the efficiency score. Thus, the details of the efficiency score assignment are hidden from the researcher.

In this paper below we present DEA Partially Perfect Objects (DEA PPO) that develops further a DEA model with a Perfect Object (DEA PO), a version of DEA aimed at

improvement of these weaknesses and simplification of the calculations. The paper is organized as follows. Section II describes DEA PO, and section III, DEA PPO. Section IV presents an example of applications. Conclusive remarks are given in section V.

II. DEA WITH A PERFECT OBJECT

Data Envelopment Analysis with a perfect Object (DEA PO) was developed in publications [5]-[7]. It was shown in these publications that if a Perfect Object is added to the group, then the efficiency score may be calculated as a ratio of the largest relative output to the smallest relative input:

$$E_k = \max_{0 \leq j \leq r} \frac{X_{0j}}{X_j} \times \max_{0 \leq i \leq s} \frac{Y_i}{Y_{0i}} = \frac{\text{maximum relative output}}{\text{minimum relative input}} \quad (7)$$

$$= \frac{X_{0j^*} Y_{i^*}}{X_{j^*} Y_{0i^*}} = \frac{\left(\frac{Y_{i^*}}{Y_{0i^*}} \right)}{\left(\frac{X_{j^*}}{X_{0j^*}} \right)}$$

where relative means expressed in terms of PO, lower indexes $i = 1..s$ and $j = 1..r$ stand for outputs and inputs, respectively, lower indexes i^* and j^* , for the maximum or minimum relative values for the output and input, respectively; lower index 0 stands for the Perfect Object. DEA PO is based on an idea of comparison with best practices. This idea is known in the DEA literature, but DEA PO extends it to (7) that provides an explicit solution.

Geometric interpretation for the DEA PO for one input and one output is given in Fig. 1 borrowed from [8]. Formula (7) may be proved in this case, based on the geometric interpretation. The proof is given in Fig. 2. In this figure, the PO is located at the point F , and an actual object - at the point G . The input - minimization efficiency score equals AB/AG , the output-minimization - DG/DK . The relative input is AG/AC , output - DG/DN . In this situation, there is no need for consideration of maximum relative output or minimum relative input, since only one input and one output are present. The formula (7) states in this case that

$$E_{IM} = AB/AG = (DG/DN)/(AG/AC) \quad (8)$$

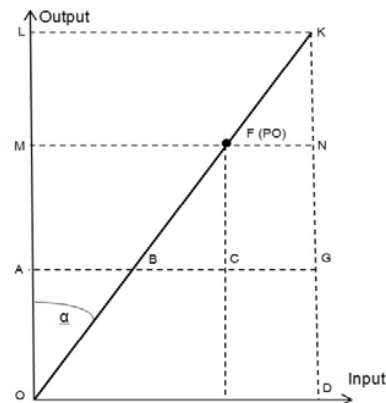
The proof is based on the proportions between the lengths of the segments cut from the sides of an angle by parallel lines:

$$\begin{aligned} (DG/DN)/(AG/AC) &= (DG/DN) \cdot (AC/AG) = \\ (AC/DN) \cdot (DG/AG) &= (MF/OM) \cdot (OA/LK) = \\ (MF/OM) \cdot (OA/OL) \cdot (OL/LK) &= \\ (OA/OL) \cdot (MF/OM) \cdot (OL/LK) &= (OA/OL) \cdot \tan(\alpha) \cdot \cot(\alpha) = \\ (OA/OL) &= AB/LK = AB/AG = E_{IM} \quad \blacksquare \end{aligned} \quad (9)$$

It may be noted that the equality of the two types of efficiency - input-minimization and output-maximization ones, $E_{IM} = E_{OM}$, follows immediately from Fig. 2 because $E_{OM} = DG/DK = AB/LK = E_{IM}$.

DEA PO improves the conventional DEA in the ability to measure absolute efficiency. If the inputs and outputs of the

III. DEA WITH PARTIALLY PERFECT OBJECTS

 HL/HG , DEA PO efficiency score, HK/HG .

Since partial efficiencies are calculated using different numbers of inputs and outputs, we assign each of them a weight proportional to the share of the total number of indicators used or its calculation.¹ The weights add up to 1. We denote partial efficiencies as

The values of w_{kl} are calculated in two steps. At the first step, we calculate the raw weights W_{kl} , then, at the second step, we adjust them to make the sum equal one. By doing so, we get $w_{kl} = W_{kl} / W$, where W_{kl} stands for the raw weights, W , for their sum, and $w_{kl} = W_{kl} / W$ for the weights scaled to sum up to 1. Adding up the partial efficiencies E_{kl} taken with their weight coefficients w_{kl} we arrive at the weighted efficiency score E_w :

This process involves all possible ratios of relative outputs to inputs in the order determined by their ranks. It results in r

²The Excel functions LARGE(*array*, *k*) and SMALL(*array*, *l*) provide a convenient tool for the calculations.

s partial efficiencies ranging from largest E_{11} to smallest E_{rs} . The last efficiency term E_{rs} is calculated based on only two indicators: the smallest relative output and the largest relative input. It is the smallest relative efficiency score.

The weighted efficiency E_w equals

$$E_w = \sum_{k=1}^s \sum_{l=1}^r w_{kl} E_{kl} \quad (12)$$

where

$$\sum_{k=1}^s \sum_{l=1}^r w_{kl} = 1 \quad (13)$$

Since the weight coefficients w_{kl} are proportional to the share of the total number of inputs and outputs that the corresponding kl -partial perfect object comprises, we have for the kl -efficiency coefficient:

$$w_{kl} = W \times \frac{(s-k-l)+(r-l-l)}{r+s} = W \times \left(1 - \frac{k+l-2}{r+s}\right) \quad (14)$$

$k = 1..s, l = 1..r,$

where r and s stand for the total number of inputs or outputs, respectively, W , for a normalizing coefficient making the sum of the weights equals to 1. From (14) it follows that

$$W = \frac{1}{\sum_{k=1}^s \sum_{l=1}^r \left(1 - \frac{k+l-2}{r+s}\right)} \quad (15)$$

The denominator of this fraction may be simplified by using the formula for the sum of an arithmetic progression:

$$\begin{aligned} \sum_{k=1}^s \sum_{l=1}^r \left(1 - \frac{k+l-2}{r+s}\right) &= \frac{1}{r+s} \times \sum_{k=1}^s \sum_{l=1}^r (r+s+2-k-l) \\ &= \frac{1}{r+s} \times \left(\sum_{k=1}^s \sum_{l=1}^r (r+s+2) - \sum_{k=1}^s \sum_{l=1}^r k - \sum_{k=1}^s \sum_{l=1}^r l \right) \\ &= \frac{rs(r+s-2)}{2(r+s)} \end{aligned} \quad (16)$$

so that

$$W = \frac{2(r+s)}{rs(r+s-2)} \quad (17)$$

The computations in (16) leading to (17) may be easily carried out by using a Computer Algebra System, such as that of the graphing calculator TI-89.

Substituting (16) into (14), we get:

$$\begin{aligned} w_{kl} &= W \times \frac{(s-k-l)+(r-l-l)}{r+s} = \frac{2(r+s-k-l+2)}{rs(r+s-2)} \quad (18) \\ &= \frac{2(r+s)}{rs(r+s-2)} \times \frac{(s-k-l)+(r-l-l)}{r+s} \end{aligned}$$

For example, in case of two inputs and two outputs, $r = 2, s = 2$, we get $w_{11} = 2(2+2-1-1+2)/(2 \cdot 2(2+2-2)) = 8/24 = 1/3$. Similar to that, we obtain that $w_{12} = w_{21} = 1/4$, and $w_{22} = 1/6$. The total is $1/3 + 1/4 + 1/4 + 1/6 = 1$. This result is used in the following section. It may also be noted that if $k + l = p + q$, then $w_{kl} = w_{pq}$.

IV. EXAMPLE OF APPLICATIONS

In this section we apply DEA PPO to firm evaluation. This problem arises when an investment portfolio is compiled. In our example, data of 10 firms, denoted as C01, C02, ..., C10, were chosen randomly from the website finance.yahoo.com. The following types of indicators were used: profitability, management effectiveness, market valuation, and volatility. For each type we chose a representative indicator as follows: profit margin (output-1), return on assets (output-2), market-value-to-revenue ratio (input-1), and the debt-to-equity ratio (input-2), respectively. The data are presented in columns 2 through 5 of the Table I. The Perfect Object is shown in the last row. It comprises maximal outputs and minimal inputs. Columns 6 through 9 contain relative outputs and inputs, respectively, calculated as the ratios of actual input or output to the corresponding value of the PO.

Column 10 contains the DEA PO efficiency score E_{11} equals the ratio of the greatest relative output to the smallest relative input. For example, for the firm C01, $E_{11} = \max(0.3150, 1.0000) / \min(4.1515, 26.6291) = 1.0000 / 4.1515 = 0.2409$. DEA PO stops in this step, but DEA PPO assigns the weight $w_{11} = 1/3 = 0.3333$ and continues. The weight was calculated in the previous section.

Column 11 presents the E_{12} efficiency scores obtained by the eliminating the smallest relative input. Depending on the particular object, it may be either relative input-1 or relative input-2, which is smaller. For the firm C01, it is $\min(4.1515, 26.6291) = 4.1515$, the relative input-1. To calculate E_{12} we apply (7) to the largest relative output-2 (1.000) and the only remaining relative input-2 equals 26.6291. Their ratio is $E_{12} = 1.0000 / 26.6291 = 0.0376$. The weight coefficient for this partial efficiency is $w_{12} = 1/4 = 0.2500$, as was calculated in the previous section.

Column 12 is similar to the column 11, except that the greatest relative output, rather than smallest relative input, is eliminated. For example, for the firm C01 the output to eliminate is $\max(0.3150, 1.0000) = 1.0000$, the relative output-2. The relative output-1 (0.3150) is retained for the calculations. Applying (7) again, we get $E_{21} = 0.3150 / \min(4.1515, 26.6291) = 0.3150 / 4.1515 = 0.0759$. The weight coefficient $w_{21} = 1/4 = 0.2500$, as was shown in previous section.

Column 13 contains the partial efficiency scores E_{22} . They are obtained by eliminating both largest relative output and

smallest relative input. For the firm C01, the indicators to eliminate are: $\max(0.3150, 1.0000) = 1.0000$, the relative output-2, and $\min(4.1515, 26.6291) = 4.1515$, the relative input-1. Remaining are the relative output-1 (0.3115) and relative input-2 (26.6291). Applying again (7), we get $E_{22} = 0.3115/26.6291 = 0.0118$. The weight coefficient $w_{22} = 1/6 = 0.1667$, as was calculated above.

Column 14 shows the weighted efficiency scores calculated by using (11). For example, for the firm C01 it is $E_w = 0.2409 \cdot 0.3333 + 0.0376 \cdot 0.2500 + 0.0759 \cdot 0.2500 + 0.0118 \cdot 0.1667 = 0.1106$. Column 15 shows the ranks of the firms based on the weighted efficiency scores. A firm with the greatest score received the rank of one. As follows from the ranking, the most efficient is the firm C08 ($E_w = 0.3748$, rank = 1), the second best is C02 ($E_w = 0.2645$, rank = 2), and the third rank is assigned to the firm C05 ($E_w = 0.1879$, rank = 3). Based on the DEA PPO analysis, these three firms may be recommended for the inclusion in the portfolio.

This case study is just an example of numerous possible applications of DEA PPO, see, for example, [8]. Simplicity of calculations allows its use at any level of preparation. Since in the DEA PPO separates the task into independent threads, the calculations may be performed by different groups of participants working independently and combining the results at the final step. The DEA PPO may be also useful as an introductory study of the conventional DEA.

V. CONCLUSIONS

This paper presents DEA with Partially Perfect Objects (DEA PPO). The DEA PPO, similar to the DEA with a Perfect Object (DEA PO), appends the group of actual objects with a virtual Perfect Object having largest outputs and smallest inputs in the group. The DEA PPO inherits the explicit solution formula of the DEA PO and its propensity to evaluate absolute efficiency. It is the case when the PO is set up based, for example, on national, international, industrial, etc. standards. However, DEA PPO uses not only the Perfect Object, but a sequence of Partially Perfect Objects, obtained by consecutive elimination of the current best relative indicator. By doing so, DEA PPO generates a series of partial efficiency scores and, finally, a weighted efficiency. The last is calculated using all possible ratios of relative outputs to relative inputs, thus preventing an object having just one large output and one small input from obtaining a high efficiency score. The computational scheme of the DEA PPO is simple and intuitive. Though, theoretically, DEA PPO uses an LP procedure, it simplifies to just calculation of a ratio. An example of applications demonstrates the technique of computations and provides a general idea of applications in different fields. Among the promising areas are teaching quantitative reasoning and introduction to DEA.

TABLE I
FIRM EVALUATION USING DEA PPO

Firm	Output-1	Output-2	Input-1	Input-2					0.3333	0.2500	0.2500	0.1667		
	Profit Margin, %	Return on Assets, %	Market Value/Revenue, r.u.	Total Debt/Equity, r.u.	Relative output-1	Relative output-2	Relative input-1	Relative input-2	E_{11}	E_{12}	E_{21}	E_{22}	E_w	Rank
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
C01	0.2892	0.1841	4.11	512.61	0.3150	1.0000	4.1515	26.6291	0.2409	0.0376	0.0759	0.0118	0.1106	4
C02	0.1317	0.1061	1.40	31.31	0.1435	0.5763	1.4141	1.6265	0.4075	0.3543	0.1082	0.0822	0.2645	2
C03	0.8058	0.0095	70.08	694.79	0.8778	0.0516	70.7879	36.0930	0.0243	0.0124	0.0014	0.0007	0.0117	9
C04	0.9180	0.0109	82.97	869.03	1.0000	0.0592	83.8081	45.1444	0.0222	0.1919	0.0013	0.0007	0.0108	10
C05	0.2011	0.0710	2.04	28.61	0.2191	0.3857	2.0606	1.4862	0.2595	0.1872	0.1474	0.1063	0.1879	3
C06	0.0987	0.0339	4.11	101.17	0.1075	0.1841	4.1515	5.2556	0.0444	0.0350	0.0259	0.0205	0.0334	8
C07	0.0098	0.0417	1.65	60.38	0.0107	0.2265	1.6667	3.1366	0.1359	0.0722	0.0064	0.0034	0.0655	5
C08	0.2158	0.1342	3.00	19.25	0.2351	0.7290	3.0303	1.0000	0.7290	0.2406	0.2351	0.0776	0.3748	1
C09	0.0436	0.0250	1.19	110.64	0.0475	0.1358	1.2020	5.7475	0.1130	0.0236	0.0395	0.0083	0.0548	6
C10	0.0380	0.0118	0.99	45.01	0.0414	0.0641	1.0000	2.3382	0.0641	0.0274	0.0414	0.0177	0.0415	7
PO	0.9180	0.1841	0.99	19.25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	

REFERENCES

- [1] Banker, R., A. Charnes, and W. Cooper. Some models for estimating technical and scale efficiency in Data Envelopment Analysis. *Management Science* 30(9): 1984; 1078–1092.
- [2] Charnes, A., W. Cooper, and E. Rhodes. Measuring the efficiency of decision-making units. *European Journal of Operational Research* 1978; 2: 429–444.
- [3] Cooper W., L. Seiford, and J. Zhu (eds.). *Handbook on data envelopment analysis*. International Series in Operations Research & Management Science, vol 164, 2nd ed. New York: Springer, 2011.

- [4] Farrell M.J. The Measurement of Production Efficiency. *Journal of the Royal Statistical Society, Series A* 1957; 120(3): 253 - 282.
- [5] Vaninsky, A. DEA with a perfect object: Analytical Solutions. *Communications in Mathematics and Applications* 2011; 2(1):1–13.
- [6] Vaninsky, A. Environmental performance of the United States Energy Sector: A DEA model with non-discretionary factors and perfect object. *Proceedings of the World Academy of Science, Engineering and Technology (WASET)* 2009; 54:139–144. Available at www.waset.org/journals/waset/v30/v30-26.pdf
- [7] Vaninsky, A. Explicit formulas for efficiency scores and weight coefficients in DEA problems with a perfect object. *International Journal of Mathematical Modeling, Simulation and Applications* 2011; 4(3): 217 –233.
- [8] Vaninsky, A. Simplified Data Envelopment Analysis: What Country Won the Olympics, and How about our CO2 Emissions? *Numeracy* 2013; 6(2) Forthcoming. Available at cholarcommons.usf.edu/numeracy.