Numeracy
Advancing Education in Quantitative Literacy

Volume 6 | Issue 2 Article 12

2013

Simplified Data Envelopment Analysis: What Country Won the Olympics, and How about our CO_2 Emissions?

Alexander Vaninsky
Hostos Community College, CUNY, avaninsky@hostos.cuny.edu

Simplified Data Envelopment Analysis: What Country Won the Olympics, and How about our CO_2 Emissions?

Abstract

This paper introduces a simplified version of Data Envelopment Analysis - 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. Examples of DEA applications discussed in this paper include the London 2012 Olympic Games and the dynamics of the United States' environmental performance. In the first example, we find a team winner and rank the teams; in the second, we analyze the dynamics of CO_2 emissions adjusted to the gross domestic product, population, and energy consumption. Adding a virtual Perfect Object – one having the greatest outputs and smallest inputs - we greatly simplify the DEA computational procedure by eliminating the Linear Programming algorithm. Simplicity of computations makes the suggested approach attractive for educational purposes, in particular, for use in Quantitative Reasoning courses.

Keywords

quantitative reasoning, operational research, multicriteria decision making, Data Envelopment Analysis, Perfect Object, Olympic Games team winner, environmental efficiency

Cover Page Footnote

Alexander Yan Vaninsky is a Professor of Mathematics at Hostos Community College of the City University of New York. He received his masters' degrees in electrical engineering from the Moscow Power Engineering Institute, in mathematics from Moscow State University, and in mathematics education from Brooklyn College; Ph.D. and D. Sc. Degrees in mathematical economics - from Moscow Finance University. His research interests include applied mathematics, technology - based mathematics education, and non - standard approaches to teaching mathematics.

Introduction

This paper introduces a tool to evaluate the efficiency and ranking of objects within a group. The objects use various resources (inputs) to produce several outputs. Because the amount of inputs and outputs is specific to each object, the question arises as to which of them produces relatively more than others. Such objects are termed *efficient.*

We examine two cases in this paper. In the first case, the "objects" are countries; in the second, the "objects" are years. The suggested tool is used to compare efficiencies of different countries in the first case, and to compare efficiencies through a succession of years in the second case.

For the first case, we look at the 2012 Summer Olympic Games in London. We determine which country was the most efficient Olympic participant, taking into consideration medals earned, the number of participating athletes, and the population of a country. In particular, the United States brought home 104 Olympic medals, of which 46 were gold, while Jamaica earned 12 medals with four of them gold. Going purely by medal count, team United States was a much more successful performer, but not necessarily the more efficient one. In terms of population, Jamaica is significantly smaller than the U.S. It has a much smaller pool from which to draw and develop athletes capable of competing at the international level. In terms of team size, the U.S. had the second largest showing at the Games with 539 athletes, compared to only 53 athletes entered by Jamaica. From an efficiency point of view, we posit that Jamaica outperformed the United States. We calculate the efficiency of both countries within the context of all Olympic Games participants to demonstrate this.

For our second case, we look at the environmental performance of the United States. Its population size and the volume of production continue to grow, increasing energy consumption, and emissions of CO_2 and other gasses. These emissions affect the integrity of the atmosphere and may strongly influence Earth's climate change. Aimed at timely intervention, an index of environmental performance can be a useful tool to monitor these processes. Such an index would objectively weigh the increase in gas emissions and energy consumption against economic output and population growth.

A mathematical tool suitable for these cases is Data Envelopment Analysis (DEA) developed in Charnes et al. (1978) and Banker et al. (1984). Its contemporary state is given in Cooper et al. (2011) and some DEA Web sites.[1, 2]

[1] "Ali Emrouznejad's Data Envelopment analysis" (2011). http://deazone.com/ (accessed April 2, 2013)

[2] "A Data Envelopment Analysis (DEA) Home Page" (1996) http://www.emp.pdx.edu/dea/homedea.html#Intro_to_DEA (accessed April 2, 2013)

DEA evaluates the performance of functionally similar objects by weighting outcomes against inputs. It uses Linear Programming (LP) to assign an *efficiency score* scaled between 0 and 1 to each object in the group.

This paper introduces a simplified version of DEA developed by Vaninsky (2009). The simplification circumvents LP by adding a virtual Perfect Object (PO). The PO receives the smallest inputs and largest outputs in the group, thereby serving as an objective benchmark for performance comparisons. An important feature of DEA PO is that efficiency scores are obtained analytically in a simple, explicit form. The derived formula requires only a moderate number of simple operations with ratios of inputs and outputs. In addition, contrary to conventional DEA, each object is processed individually, independently of all other objects in the group.

The theory of DEA PO was further developed by Vaninsky (2011a,b), in line with various suggestions to modify the DEA in order to include “best practices” in the analysis. The following publications are some that include the best practices approach: Cook and Zhu (2006), Golany (1988), Thanassoulis and Dyson (1992), and Zhu (1996). However, all of these publications remain within the framework of the LP algorithms or other optimization procedures; not one suggests an explicit analytic solution to the corresponding DEA problem.

As with any procedure of multicriteria comparison, DEA PO is not perfect. We discuss some of its shortcomings and introduce a means for its improvement.

Lastly, the suggested approach is suitable for teaching undergraduate courses in Quantitative Reasoning (QR), Consumer Mathematics, or Introduction to Operations Research. Select elements may be appropriate for enrichment courses in junior high or secondary school. Besides critical thinking, the main QR skills here involve ratios, sums, and weighted averages.

Finding a Team Winner of the Olympic Games

While the Olympics typically recognize individual achievements, tracking the overall performance of nation states draws great public interest and positively contributes to a sense of national pride. At the conclusion of the Olympics, when the final medal count is in, two criteria determine which nation is the overall winner of the Games: (1) the number of gold medals won and (2) the total number of all medals won by that country. Going by these two criteria, the 2012 London Olympic Games crowned the United States the overall winner with 46 gold medals and 104 total medals. China trailed the U.S. with 38 gold and 87 total medals.[3] It turns out that such criteria are advantageous only to large countries

[3] “London 2012.” The London Organizsing Committee of the Olympic Games and Paralympic Games Ltd (LOCOG). http://www.london2012.com/medals/medal-count/ (accessed April 3, 2013).

DOI: http://dx.doi.org/10.5038/1936-4660.6.2.12

capable of sending large teams to the Olympics. These statistics inherently lack a degree of meaningfulness due to the differing population sizes of nations present at the Games and the size of their Olympic teams. The playing field is not level when the overall winner is determined by the size of the medal haul alone. For example, the United States has a total population of 314.1 million people and sent a 539-member team to the Games. China has 1,347.4 million people and sent 385 athletes to compete. Jamaica has only 2.7 million citizens and had 53 athletes; New Zealand has a population of 4.4 million and sent 115 athletes to the 2012 Games. We, therefore, suggest a different approach for selecting the overall winning country by adjusting the numbers of gold and total medals for country population and size of their Olympic team.

To determine which nation is most effective in winning medals, we find an Olympic team that is: (1) able to win a sufficient number of medals with a high percentage of gold medals among these, (2) moderate in size, and (3) comparable in the rank of its size to the rank of the country's population. The underlying hypothesis is that sports talent is uniformly distributed around the world, and it is a matter of national sports policy to: (1) create the right conditions to prepare athletes that are able to meet the Olympics qualifying standards,[4] and (2) form a team of optimal size.

By doing so, we deliberately ignore some fundamental factors, such as nation’s wealth and health indicators, the share of GDP it directs towards sports, its culture, environmental conditions that may or may not be favorable for its sports industry, etc. We do this for several reasons. Firstly, our main goal in this paper is to introduce DEA PO as an evaluation and decision-making tool. We want to use as simple model as possible to focus on the method itself. Secondly, the fundamental factors may be used in the next step of analysis aimed at revealing their impact on the efficiency.

To find a national team winner of the XXX Olympic Games in London we apply DEA PO relying on information provided by the London 2012 Web site.[5] Population statistics were collected from Wikipedia.[6] A total of 10,961 athletes representing 205 countries participated in the 2012 Games. In total, they were awarded 958 medals: 304 gold, 297 silver, and 357 bronze. The averages were as follows: 4.67 medals per country and 11.44 athletes per medal; only 87 countries claimed any kind of medal.

[4] “Athletics at the 2012 Summer Olympics – Qualification.” http://en.wikipedia.org/wiki/Athletics_at_the_2012_Summer_Olympics_-_Qualification (accessed April 3, 2013).

[5] http://www.london2012.com/medals/medal-count/ (See Footnote 3).

[6] “List of countries by population.” http://en.wikipedia.org/wiki/List_of_countries_by_population (accessed April 3, 2013).

To simplify our analysis, we limited our considerations to 35 countries that won at least two gold medals and at least four medals in total. Countries, their medals, team sizes, and their populations are shown in Table 1. The countries are sorted by total medal count.

TABLE 1.
COUNTRIES, MEDALS, TEAM SIZES, AND POPULATION [1)]

Country	Abbr.	Gold medals Output-1[2)]	Total medals Output-2[2)]	Team size Input-1[2)]	Population, mln. Input-2[2)]
(1)	(2)	(3)	(4)	(5)	(6)
United States of America	USA	**46**	**104**	539	314.1
Peoples Republic of China	Chi	38	87	385	1347.4
Russian Federation	Rus	24	82	441	143.1
Great Britain	GBr	29	65	558	62.3
Germany	Ger	11	44	399	81.9
Japan	Jap	7	38	305	127.5
Australia	Aus	7	35	414	22.7
France	Fra	11	34	340	65.4
Republic of Korea	SKr	13	28	289	50.0
Italy	Ita	8	28	258	60.8
Netherlands	Nth	6	20	182	16.7
Ukraine	Ukr	6	20	238	45.6
Hungary	Hun	8	17	267	10.0
Brazil	Bra	3	17	159	192.4
Cuba	Cub	5	14	110	11.2
Kazakhstan	Kaz	7	13	173	16.8
New Zeland	NZ	5	13	115	4.4
Belarus	Bel	3	13	195	9.5
Jamaica	Jam	4	12	53	**2.7**
Islamic Republic of Iran	Ira	4	12	50	75.2
Kenya	Ken	2	11	50	42.7
Czech Republic	Cze	4	10	133	10.5
Poland	Pol	2	10	220	38.5
Romnania	Rom	2	9	53	19.0
Denmark	Den	2	9	115	5.6
Azerbaijan	Aze	2	9	105	9.2
Spain	Spa	3	8	293	46.2
Ethiopia	Eth	3	7	**35**	84.3
Democratic Peoples Republic of Korea	NKr	4	6	109	24.6
Croatia	Cro	3	6	56	4.3
South Africa	Saf	3	6	133	50.6
Lithuania	Lth	2	5	62	3.2
Turkey	Tur	2	5	114	74.7
Norway	Nor	2	4	65	5.0
Switzerland	Swi	2	4	106	8.0
Perfect Object	**PO**	**46**	**104**	**35**	**2.7**

Notes:
[1)] Sorted by total medals (Column 4).
[2)] Indicators forming the Perfect Object are shown in bold.

In this example, the Perfect Object is a hypothetical Olympic team that: (1) has won a maximum number of both gold and total medals, (2) represents a country with the smallest population size, and (3) has the least number of athletes on their Olympic team. The Perfect Object receives the highest possible DEA

efficiency score of 1, while the rest of the teams receive smaller scores, ranging from 0 to 1. A national team that receives the highest score is declared the overall winner; all other teams are ranked by their efficiency scores. Mathematically, a PO defines a frontier of opportunities in the four-dimensional space *Gold medals - Total medals - Team size - Population,* with all actual teams located on one side of the frontier. Efficiency, as measured by DEA, is proximity to the frontier: the closer a country is to the frontier, the higher the efficiency score. DEA PO suggests the following simple formula:

$$E = \frac{maximum\ relative\ output}{minimum\ relative\ input}, \tag{1}$$

where *relative* implies expression in terms of the Perfect Object.

Since all actual outputs and inputs are positive, with outputs not exceeding the PO values, and actual inputs being no less than those of the PO, the numerator is at most 1, whereas the denominator is at least 1. Thus the value of E in Equation 1 is between 0 and 1, as required by DEA. More detail is given in the Appendix.

The Perfect Object (PO) is formed by taking the largest outputs and smallest inputs. Countries forming the PO in this example are as follows: USA - Gold and total medals (46 and 104, respectively), Ethiopia - Team size (35 athletes), and Jamaica (population of 2.7 million people). Table 2 shows the relative inputs and outputs obtained by taking the ratios of actual values to the corresponding values of the PO, and efficiency calculations.

For example, for the United States, the relative output-1 (gold medals) is 46/46 = 1.000, relative output-2 (total medals) is 104/104 = 1.0000, relative input-1 (team size) is 539/35 = 15.4000, and relative input-2 (population size) is 314.1/2.7 = 116.1002. These numbers are shown in columns 2, 3, 5, and 6, respectively. To calculate efficiency, we need the maximum value of relative outputs and the minimum value of relative inputs. For example, for the USA, the maximum relative output is *max* (1.0000, 1.0000) = 1.000, and minimum relative input is *min*(15.4000, 116.1002) = 15.4000, as shown in columns 4 and 7. To calculate the efficiency score, we use Equation 1 and take the ratio of maximum relative output to the minimum relative input. For USA team, the result is 1.0000/15.4000 = 0.0649, as shown in column 8. Column 9 presents efficiency ranks with rank number 1 assigned to the country with the highest efficiency score.

As follows from Table 2, the overall winner is Jamaica with an efficiency score of 0.1154. It is followed by Iran (0.0808), New Zealand (0.0763) and China (0.0760). Jamaica is the smallest by population: only 2.7 million, and its team of

53 is very close to an average of 53.5. The Jamaican team was very successful in track and field, and won 12 Olympic medals with 4 gold medals between them.

The Jamaican team won one Olympic medal for every 4.41 team members, while the average is one medal for every 11.44 team members. The ratio of population (2.706 million) to team size is just 51.1 for Jamaica, while for the sample it is 433.5.

Iran was second in efficiency rankings. It is very close to Jamaica by gold and total medals, as well as team size: 4 gold and 12 total medals, 50 member team. However, at 75.15 million people, Iran's population is much larger. Based on this observation, one may speculate that a team of 50–53 is optimal for countries with a limited Olympic budget.

TABLE 2.
EFFICIENCY

Country	Relative Output-1	Relative Output-2	Max Relative Output	Relative Input-1	Relative Input-2	Min Relative Input	Efficiency[1)]	Efficiency rank
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
USA	1.0000	1.0000	1.0000	15.4000	116.1002	15.4000	0.0649	7
Chi	0.8261	0.8365	0.8365	11.0000	497.9439	11.0000	0.0760	4
Rus	0.5217	0.7885	0.7885	12.6000	52.8921	12.6000	0.0626	8
GBr	0.6304	0.6250	0.6304	15.9429	23.0103	15.9429	0.0395	17
Ger	0.2391	0.4231	0.4231	11.4000	30.2529	11.4000	0.0371	18
Jap	0.1522	0.3654	0.3654	8.7143	47.1316	8.7143	0.0419	12
Aus	0.1522	0.3365	0.3365	11.8286	8.3886	8.3886	0.0401	16
Fra	0.2391	0.3269	0.3269	9.7143	24.1516	9.7143	0.0337	24
SKr	0.2826	0.2692	0.2826	8.2571	18.4803	8.2571	0.0342	23
Ita	0.1739	0.2692	0.2692	7.3714	22.4777	7.3714	0.0365	20
Nth	0.1304	0.1923	0.1923	5.2000	6.1862	5.2000	0.0370	19
Ukr	0.1304	0.1923	0.1923	6.8000	16.8399	6.8000	0.0283	27
Hun	0.1739	0.1635	0.1739	7.6286	3.6817	3.6817	0.0472	10
Bra	0.0652	0.1635	0.1635	4.5429	71.0971	4.5429	0.0360	21
Cub	0.1087	0.1346	0.1346	3.1429	4.1569	3.1429	0.0428	11
Kaz	0.1522	0.1250	0.1522	4.9429	6.2000	4.9429	0.0308	25
NZ	0.1087	0.1250	0.1250	3.2857	1.6389	1.6389	0.0763	3
Bel	0.0652	0.1250	0.1250	5.5714	3.4952	3.4952	0.0358	22
Jam	0.0870	0.1154	0.1154	1.5143	1.0000	1.0000	0.1154	1
Ira	0.0870	0.1154	0.1154	1.4286	27.7733	1.4286	0.0808	2
Ken	0.0435	0.1058	0.1058	1.4286	15.7989	1.4286	0.0740	5
Cze	0.0870	0.0962	0.0962	3.8000	3.8821	3.8000	0.0253	29
Pol	0.0435	0.0962	0.0962	6.2857	14.2289	6.2857	0.0153	32
Rom	0.0435	0.0865	0.0865	1.5143	7.0378	1.5143	0.0571	9
Den	0.0435	0.0865	0.0865	3.2857	2.0640	2.0640	0.0419	13
Aze	0.0435	0.0865	0.0865	3.0000	3.4130	3.0000	0.0288	26
Spa	0.0652	0.0769	0.0769	8.3714	17.0606	8.3714	0.0092	35
Eth	0.0652	0.0673	0.0673	1.0000	31.1627	1.0000	0.0673	6
NKr	0.0870	0.0577	0.0870	3.1143	9.0745	3.1143	0.0279	28
Cro	0.0652	0.0577	0.0652	1.6000	1.5857	1.5857	0.0411	14
Saf	0.0652	0.0577	0.0652	3.8000	18.6955	3.8000	0.0172	31
Lth	0.0435	0.0481	0.0481	1.7714	1.1781	1.1781	0.0408	15
Tur	0.0435	0.0481	0.0481	3.2571	27.6161	3.2571	0.0148	34
Nor	0.0435	0.0385	0.0435	1.8571	1.8585	1.8571	0.0234	30
Swi	0.0435	0.0385	0.0435	3.0286	2.9391	2.9391	0.0148	33
PO	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	

Note. [1)] Efficiency = Max relative output/Min relative input

The efficiency ratings of China (385-member team, ranked 4th), United States (539-member team , ranked 7th), Russia (441-member team, ranked 8th), and Great Britain (558-member team, ranked 17th) are much lower. These countries experienced "diminishing returns to scale" effect and have not managed to win enough medals to justify their large teams.

Partial Perfect Object, Partial Efficiency, and Weighted Efficiency

The results of the preceding section shed light on the nature and properties of DEA PO that are important for its practical applications. Firstly, DEA PO, similar to conventional DEA, measures only the relative efficiency. The obtained efficiency score determines the position of a particular country within the sample, but by no means does it reveal its entire performance during the Olympic Games. Another observation is that DEA PO evaluates objects by their best ratios only. It may assign a high efficiency score for just one high ratio that appears in Equation 1, no matter how large or small all of the other indicators are.

For example, Jamaica received its highest rank based on the relatively large total number of medals (12) and a small population (2.7 million.). The number of gold medals (4 medals) and the team size (53 athletes) were completely ignored. This means that Jamaica could send a team of any size, say as large as that of Great Britain (558 athletes) and win just two gold medals—to be included in the list for analysis—and still get the same highest rank. This observation, at least potentially, may make us less confident in the obtained results. Some practitioners may even consider this a flaw. However, more detailed consideration reveals that it is simply the nature of the DEA PO methodology—to evaluate objects by their best ratios. The conventional DEA also evaluates efficiency based on larger outputs and smaller inputs.[7] This fact has been mentioned in the DEA literature, and different extensions of the standard DEA procedures have been suggested to improve the situation if this property is not acceptable (Andersen and Petersen 1993; Daneshian et al. 2005).

The DEA PO inherits this property from conventional DEA and uses the ratio of the largest relative output to the smallest relative input only, i.e., the highest possible ratio. In this section we extend DEA PO to incorporate all of the ratios,

[7] DEA, theoretically, 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. But the DEA optimal solution typically assigns the non-zero weights only to some of them, thus making it possible to increase other inputs or decrease other outputs arbitrarily without any change in the efficiency score. See the Appendix section for more detail.

thus eliminating the undesirable property of dealing with the best ratio only. To do that, we introduce a collection of partial PO's and corresponding partial efficiencies. Taken together, they allow for a more comprehensive evaluation of the Olympic team efficiency and a more justifiable ranking of the teams. In ranking the Olympic teams, Partial PO's are generated by consecutive elimination of (1) currently best input or output, and (2) two currently best indicators—two smaller inputs, two larger outputs, or one smallest input and one largest output, correspondingly. Every time, we apply Equation 1 to evaluate *partial efficiency* using the remaining indicators.

Because 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.[8] The weights add up to 1. We denote partial efficiencies as

$$E_{kl} = \frac{k^{th} \text{ largest output}}{l^{th} \text{ smallest input}}, \tag{2}$$

and refer to this modification of the DEA PO as DEA wth Partial PO's (DEA PPO). In this notation, the DEA PO efficiency provided by Equation 1 becomes the partial efficiency E_{11}.

The process of forming the weight coefficients is as follows. For an object that has r inputs and s outputs, a partial efficiency score E_{kl} is obtained by eliminating $(k-1)$ outputs and $(l-1)$ inputs. This means that only only $(r + s - k - l + 2)$ indicators out of the total number of $(r + s)$ remain. The smaller number of participating indicators leads to the smaller weight that the E_{kl} receives with regards to the highest efficiency score of E_{11} that is calculated using all inputs and outputs. In this paper we suggest that the weight coefficients w_{kl} assigned to the efficiency scores E_{kl} be proportional to the share of the total number of inputs and outputs that each particular E_{kl} incorporates.[9] The values of w_{kl} are calculated in two steps. At the first step, we calculate the raw weights W_{kl}, and, at the second, we adjust them to make their sum equal to 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 W_{kl} taken with their weight coefficients w_{kl} we arrive at the weighted efficiency score E_w:

$$E_w = w_{11}\cdot E_{11} + w_{12}\cdot E_{12} + \ldots w_{rs}\cdot E_{rs}. \tag{3}$$

[8] This way of assigning weights is not unique; we use it in this paper as a reasonable approach.

[9] The Excel functions LARGE(*array*, *k*) and SMALL(*array, l*) offer a convenient tool for the calculations.

This process results in $r \times s$ partial efficiencies ranging from largest E_{11} to smallest E_{rs}. The last efficiency term E_{rs} contains only two indicators: one smallest output and one largest input. It furnishes the lowest efficiency score by using the smallest possible ratio of relative output to relative input. The appendix provides more detail.

As an example, consider the formation of the weight coefficients and generation of the PPO's for the Jamaican team. For this team, E_{11} = 0.1154, the efficiency score calculated by using Equation 1. Eliminate the minimum input *Team size* first. As follows from Table 2, the corresponding partial efficiency is E_{12} = 0.1154 ÷ 1.5143 = 0.0762. Then eliminate the maximum output *Total medals.* The partial efficiency is E_{21} = 0.0870 ÷ 1.0000 = 0.0870. Finally, eliminate both the maximum relative output and relative minimum input: E_{22} = 0.0870 ÷ 1.5143 = 0.0574.

In this process the DEA PO efficiency score E_{11} uses four indicators; the partial efficiencies E_{12} and E_{21} use three indicators; and E_{22} uses two. The raw weights that are proportional to the share of the total number of inputs and outputs but do not add up to 1 are as follows: W_{11} = 4/4 = 1.000, W_{12} = W_{21} = ¾ = 0.75, and W_{22} = 2/4 = 0.5000. The total of the raw weights is W = 1.0000 + 0.75 + 0.75 + 0.5000 = 3.0000. To scale the raw weights, divide each of them by their total W = 3.000, thus making them adding up to 1: w_{11} = 1.0000 ÷ 3.0000 = 0.3333, w_{12} = w_{21} = 0.75 ÷ 3.0000 = 0.2500, and w_{22} = *0.5000 ÷ 3.000*0 = 0.1667. Check: 0.3333 + 0.2500 + 0.2500 + 0.1667 = 1.0000, as required. After doing this, the weighted efficiency score E_w is calculated as the weighted average of the partial efficiencies: E_w = 0.1154 × 0.3333 + 0.0762 × 0.25 + 0.0870 × 0.25 + 0.0574 × 0.1667 = 0.0888.[10]

By doing this for all other countries in the sample, we get the results shown in Table 3 and Figures 1 and 2. Any partial efficiency E_{kl} is smaller than the DEA PO efficiency E_{11} because the latter uses all inputs and outputs and chooses the best two. As the largest value of E_{11} becomes just one of the weighted components, the weighted efficiency E_w is less than E_{11}. In view of this observation, it should be stressed once again that the DEA efficiency score should be treated as an indicator of relative performance only. As follows from Table 3 and the graphs, Jamaica, measured by the weighted efficiency, retains its leading position. Iran moves to the fourth position, while New Zealand moves into second. China and the United States improve their ranks to 3rd and 5th, respectively.

The following question arises: which efficiency scores are preferable? The answer depends on the objectives and the type of analysis. In the case of the

[10] It is shown in the Appendix that in a general case of r inputs and s outputs, $W_{kl} = (r+s+2-k-l)\,/\,(r+s)$; $W = rs(r+s+2)/2(r+s)$; $w_{kl} = W_{kl}/W = 2(r+s-k-l+2)/rs(r+s+2)$; $k = 1, ..., s$; $l = 1, …, r$.

Olympic Games, it seems that weighted efficiency gives a more objective evaluation, while taking the average of both may be even more preferable. While using the average of the DEA PO efficiency and the weighted efficiency, Jamaica is again the team winner, followed by New Zealand, Iran, China, and Kenya, with the US team ranked 7th. Interestingly, Jamaica retains top position with any kind of calculation. (Table 3 provides more detail.)

TABLE 3.
WEIGHTED AND AVERAGE EFFICIENCY

	E_{11} w_{11}= 0.3333	E_{12} w_{12} = 0.2500	E_{21} w_{21} = 0.2500	E_{22} w_{22} = 0.1667	E_w	Weighted efficiency rank	Average efficiency[1)]	Average efficiency rank
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
USA	0.0649	0.0086	0.0649	0.0086	0.0415	5	0.0532	7
Chi	0.0760	0.0017	0.0751	0.0017	0.0448	3	0.0604	4
Rus	0.0626	0.0149	0.0414	0.0098	0.0366	8	0.0496	8
GBr	0.0395	0.0273	0.0392	0.0271	0.0343	12	0.0369	14
Ger	0.0371	0.0139	0.0210	0.0079	0.0224	22	0.0298	21
Jap	0.0419	0.0077	0.0175	0.0032	0.0208	27	0.0314	18
Aus	0.0400	0.0285	0.0181	0.0129	0.0271	17	0.0336	16
Fra	0.0337	0.0135	0.0246	0.0099	0.0224	23	0.0280	24
SKr	0.0342	0.0153	0.0326	0.0145	0.0258	19	0.0300	19
Ita	0.0365	0.0120	0.0236	0.0077	0.0223	24	0.0294	22
Nth	0.0370	0.0311	0.0251	0.0211	0.0299	15	0.0334	17
Ukr	0.0283	0.0114	0.0192	0.0077	0.0184	28	0.0233	28
Hun	0.0470	0.0228	0.0441	0.0214	0.0360	9	0.0415	10
Bra	0.0360	0.0023	0.0144	0.0009	0.0163	30	0.0261	25
Cub	0.0428	0.0325	0.0346	0.0262	0.0354	10	0.0391	12
Kaz	0.0308	0.0245	0.0253	0.0201	0.0260	18	0.0284	23
NZ	0.0767	0.0380	0.0667	0.0331	0.0573	2	0.0670	2
Bel	0.0355	0.0224	0.0185	0.0117	0.0240	21	0.0298	20
Jam	0.1154	0.0762	0.0870	0.0574	0.0888	1	0.1021	1
Ira	0.0808	0.0041	0.0609	0.0031	0.0437	4	0.0622	3
Ken	0.0740	0.0067	0.0304	0.0027	0.0344	11	0.0542	5
Cze	0.0253	0.0247	0.0229	0.0224	0.0241	20	0.0247	27
Pol	0.0153	0.0067	0.0069	0.0030	0.0090	33	0.0122	33
Rom	0.0571	0.0123	0.0287	0.0062	0.0303	14	0.0437	9
Den	0.0417	0.0263	0.0210	0.0132	0.0279	16	0.0348	15
Aze	0.0288	0.0254	0.0145	0.0128	0.0217	26	0.0253	26
Spa	0.0092	0.0045	0.0078	0.0038	0.0068	35	0.0080	35
Eth	0.0673	0.0022	0.0652	0.0021	0.0396	6	0.0535	6
NKr	0.0279	0.0095	0.0185	0.0063	0.0174	29	0.0227	30
Cro	0.0410	0.0408	0.0362	0.0361	0.0389	7	0.0399	11
Saf	0.0172	0.0035	0.0152	0.0031	0.0109	32	0.0140	32
Lth	0.0406	0.0271	0.0367	0.0245	0.0336	13	0.0371	13
Tur	0.0148	0.0017	0.0133	0.0016	0.0090	34	0.0119	34
Nor	0.0235	0.0234	0.0208	0.0207	0.0223	25	0.0229	29
Swi	0.0147	0.0144	0.0130	0.0127	0.0138	31	0.0143	31

Note. [1)] Average efficiency = (E_{11} + E_w)/2

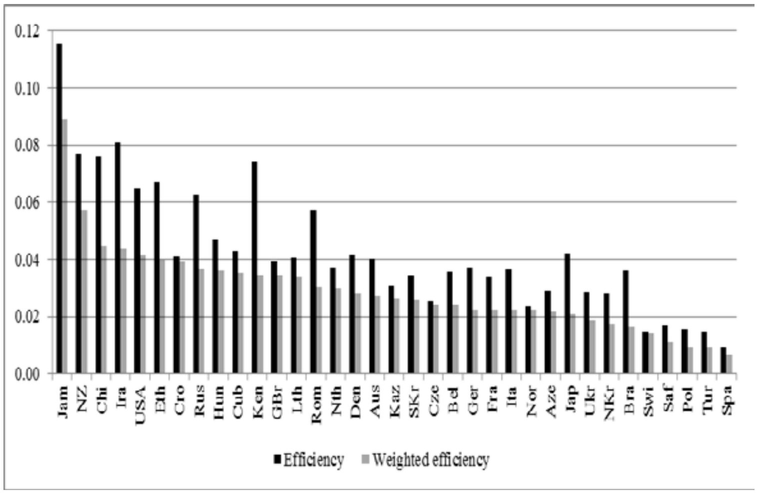

Figure 1. DEA PO efficiency and weighted efficiency

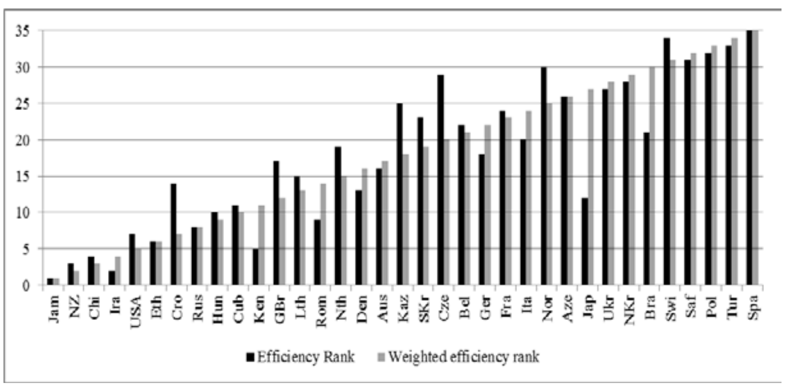

Figure 2. Ranks of DEA PO efficiency and weighted efficiency. Lower rank corresponds to greater efficiency.

Dynamics of the United States' Environmental Efficiency

Since the end of the 20th century, environmental problems have come to the forefront of national and international concerns. In large part, the increasing concerns have accompanied the ramping up of the economies and industrial ouput of underdeveloped and developing nations. Increasingly, the public is aware that the amount of CO_2 and other gas emissions have steadily increased, imposing far-reaching effects on the atmosphere, global warming and climate change. A variety of the problems related to these issues have been investigated using DEA (Zhou et

al. 2008). From the DEA point of view, analysis of environmental performance is an example where the Perfect Object appears in a natural way. It represents either international standards regarding the gas emissions, or actually observed best practices. Environmental performance is evaluated as the outcome of the interplay of the two groups of factors acting in opposite directions. The first group comprises positive factors of economic development: an increase in output, and population growth. The second group contains negative factors, actually resulting from the first group: higher energy consumption and greater emissions. In this section we apply DEA PO to investigate the environmental efficiency of the United Sates for the period of 1980–2010. We examine the environmental efficiency in dynamics, thus answering the question: How are we doing with our CO_2 Emissions?

In terms of DEA, the CO_2 emissions are seen as an output because they constitute the physical outcome of economic and technological processes. But it is an *undesirable output* and, as such, should be decreased rather than increased, contrary to conventional outputs. Undesirable outputs were introduced in Färe et al. (1989), and Seiford and Zhu (2002) suggested a corresponding DEA model. In the case of undesired outputs, DEA uses *translated outputs* that are the complements of actual outputs to some large values V_i, specific for each undesired output.[11] DEA assigns higher efficiency scores for greater translated outputs and, thus, for smaller undesired ones. In this paper, we use energy-related CO_2 emissions as undesirable output, in order to evaluate the extent to which their increase is justified by the growth of the gross domestic product (GDP) and population. The value of a parameter V was taken as

$$V = (maximum\ CO_2\ output) + (minimum\ CO_2\ output). \tag{4}$$

This choice of V swaps maximum and minimum values and guarantees positiveness of the translated output. We apply DEA PO and DEA PPO to statistical data of the United States for years 1980 through 2010, available on the Web site of the U.S. Energy Information Agency.[12] In addition to the CO_2 emissions as an undesirable output, we use GDP as a conventional output and population and energy consumption as inputs.

When evaluating the dynamics of environmental efficiency, we use three efficiency measures. The first two are the same as in the previous section: the DEA PO efficiency and the weighted efficiency. An additional indicator used in

---

[11] Impact of a specific choice of V_i values on efficiency scores has not been fully investigated in the DEA literature yet.

[12] "eia: Independent Statistics & Analysis, U.S. Energy Information Administration." www.eia.gov (accessed April 3, 2013).

this section is the efficiency score measured by the ratio of minimum output to maximum input. This smallest efficiency ratio allows for the evaluation of the worst dimension of environmental efficiency; its dynamics are important in the decision making where strategies for the improvement of environmental performance are concerned.

Table 4.
Inputs and outputs for environmental efficiency investigation

Year	GDP	CO_2 emissions[1)]	Translated CO_2 emissions[2)]	Energy consumption	Population [3)]
	Bln. $2005	Mln. metric tons	Mln. metric tons	Quadrillion Btu	Million persons
	Output-1	Undesired Output	Output-2	Input-1	Input-2
(1)	(2)	(3)	(4)	(5)	(6)
1980	5834	4770	5633	78	227
1981	5982	4642	5761	76	229
1982	5866	4406	5997	73	232
1983	6131	4383	6020	73	234
1984	6572	4613	5790	77	236
1985	6843	4600	5803	76	238
1986	7081	4608	5795	77	240
1987	7307	4764	5639	79	242
1988	7607	4982	5421	83	244
1989	7879	5067	5336	85	247
1990	8027	5039	5364	84	250
1991	8008	4996	5407	84	253
1992	8280	5093	5310	86	257
1993	8516	5185	5218	87	260
1994	8863	5258	5145	89	263
1995	9086	5314	5089	91	266
1996	9426	5501	4902	94	269
1997	9846	5575	4828	95	273
1998	10275	5622	4781	95	276
1999	10771	5682	4721	97	279
2000	11216	5867	4536	99	282
2001	11338	5759	4644	96	285
2002	11543	5806	4597	98	288
2003	11836	5857	4546	98	290
2004	12247	5975	4428	100	293
2005	12623	5997	4406	100	296
2006	12959	5919	4484	100	298
2007	13206	6020	4383	101	301
2008	13162	5838	4565	99	304
2009	12703	5429	4974	95	307
2010	13088	5612	4791	98	309
PO	13206		6020	73	227

Notes
[1)] Energy- related CO_2 equivalent emissions
[2)] Translated output = W - actual emissions, where W = 10,403
[3)] The resident population of the 50 states and the District of Columbia estimated for July 1 of each year.

Table 5.
Relative inputs and outputs and efficiencies

Year	Relative output-1[1)]	Relative output-2[1)]	Relative input-1[1)]	Relative input-2[1)]	E_{11}[2)]	E_{12}[2)]	E_{21}[2)]	E_{22}[2,3)]	E_w[4)]	Average efficiency[5)]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1980	0.4418	0.9357	1.0698	1.0000	0.9357	0.8746	0.4418	0.4129	0.7098	0.6658
1981	0.4530	0.9570	1.0430	1.0099	0.9476	0.9176	0.4485	0.4343	0.7298	0.6939
1982	0.4442	0.9962	1.0018	1.0195	0.9944	0.9771	0.4434	0.4357	0.7592	0.7240
1983	0.4642	1.0000	1.0000	1.0289	1.0000	0.9719	0.4642	0.4512	0.7676	0.7302
1984	0.4976	0.9618	1.0502	1.0378	0.9267	0.9158	0.4795	0.4738	0.7367	0.7088
1985	0.5182	0.9640	1.0469	1.0471	0.9208	0.9206	0.4950	0.4949	0.7433	0.7196
1986	0.5361	0.9626	1.0504	1.0568	0.9165	0.9109	0.5104	0.5073	0.7453	0.7212
1987	0.5533	0.9367	1.0834	1.0663	0.8785	0.8646	0.5189	0.5107	0.7238	0.6997
1988	0.5760	0.9005	1.1335	1.0760	0.8369	0.7945	0.5353	0.5082	0.6961	0.6663
1989	0.5966	0.8864	1.1619	1.0862	0.8160	0.7629	0.5493	0.5135	0.6856	0.6540
1990	0.6078	0.8910	1.1578	1.0986	0.8111	0.7696	0.5533	0.5250	0.6886	0.6610
1991	0.6064	0.8982	1.1572	1.1134	0.8067	0.7762	0.5447	0.5240	0.6865	0.6622
1992	0.6270	0.8821	1.1756	1.1289	0.7813	0.7503	0.5554	0.5333	0.6758	0.6531
1993	0.6449	0.8668	1.1981	1.1439	0.7577	0.7235	0.5637	0.5382	0.6641	0.6419
1994	0.6711	0.8547	1.2209	1.1580	0.7380	0.7000	0.5796	0.5497	0.6575	0.6357
1995	0.6880	0.8453	1.2475	1.1719	0.7214	0.6776	0.5871	0.5515	0.6486	0.6259
1996	0.7137	0.8143	1.2885	1.1856	0.6868	0.6320	0.6020	0.5539	0.6298	0.6052
1997	0.7455	0.8020	1.2964	1.1999	0.6684	0.6186	0.6213	0.5751	0.6286	0.6074
1998	0.7780	0.7942	1.3021	1.2140	0.6542	0.6099	0.6409	0.5975	0.6303	0.6126
1999	0.8156	0.7842	1.3245	1.2280	0.6641	0.6157	0.6386	0.5921	0.6336	0.6138
2000	0.8493	0.7535	1.3542	1.2418	0.6840	0.6272	0.6068	0.5564	0.6292	0.6043
2001	0.8585	0.7714	1.3179	1.2541	0.6845	0.6514	0.6151	0.5853	0.6424	0.6264
2002	0.8741	0.7636	1.3381	1.2658	0.6905	0.6532	0.6033	0.5707	0.6394	0.6211
2003	0.8963	0.7551	1.3427	1.2767	0.7020	0.6675	0.5915	0.5624	0.6425	0.6241
2004	0.9273	0.7355	1.3726	1.2886	0.7196	0.6756	0.5708	0.5359	0.6408	0.6174
2005	0.9558	0.7319	1.3743	1.3005	0.7349	0.6955	0.5628	0.5326	0.6483	0.6255
2006	0.9812	0.7449	1.3653	1.3131	0.7472	0.7187	0.5672	0.5455	0.6615	0.6419
2007	1.0000	0.7281	1.3882	1.3257	0.7543	0.7204	0.5492	0.5245	0.6562	0.6337
2008	0.9966	0.7583	1.3605	1.3383	0.7447	0.7326	0.5666	0.5574	0.6659	0.6520
2009	0.9619	0.8262	1.2959	1.3501	0.7423	0.7125	0.6376	0.6120	0.6869	0.6705
2010	0.9910	0.7958	1.3392	1.3613	0.7400	0.7280	0.5943	0.5846	0.6747	0.6624

Notes.
[1)] In terms of PO.
[2)] E_{kl} equals to the ratio of the k-th largest output to the l-th smallest input
[3)] The smallest environmental ratio.
[4)] Weighted efficiency: $w_{11} = 0.3333$, $w_{12} = 0.2500$, $w_{21} = 0.2500$, $w_{22} = 0.1667$.
[5)] The average of E_{11}, E_{22}, and E_w.

relative inputs and outputs measured in terms of the PO, as well as different types of environmental efficiencies. In this table, E_{11} is the DEA PO efficiency given by the Equation 1. Efficiency scores E_{12}, E_{21}, and E_{22} are calculated by using Equation 2. Efficiency E_{22} uses the ratio of smallest output to largest input, thus representing the worst environmental indicator. The weighted efficiency E_w is calculated by using the weights obtained above: $w_{11} = 0.3333$, $w_{12} = w_{21} = 0.2500$, and $w_{22} = 0.1667$.

Figures 3 and 4 depict graphically the entries in the Table 4 and ratios calculated from them, respectively. The graphs reveal increasing trends in GDP and population, and show that the dynamics of energy consumption and energy-related CO_2 emissions are mixed.

DOI: http://dx.doi.org/10.5038/1936-4660.6.2.12

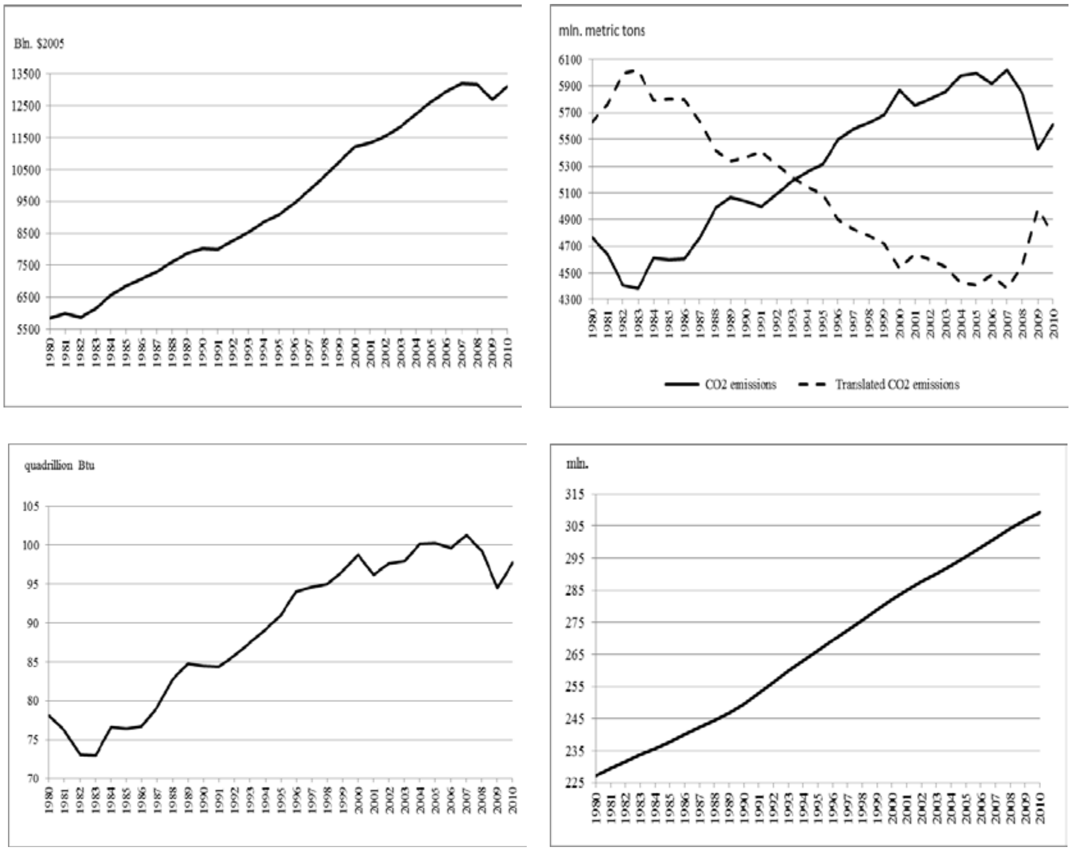

Figure 3. Factors of environmental efficiency. Graphs of 1980–2010 data in Table 4: GDP in billions of \$2005 (upper left), CO_2 emissions and translated CO_2 emissions in millions of metric tons (upper right), energy consumption in quadrillions of Btu (lower left) and population in millions of people (lower right).

Figure 4 presents the environmental ratios calculated based on the Table 4. The normalized ratios, recalculated in terms of the Perfect Object, are given in Table 5. The interplay of the ratios results in the dynamics of efficiency scores shown in the columns 6 - 11 of Table 2 and in Figure 5. It follows from the figure that the dynamics of the DEA PO efficiency (E_{11}, the largest) and of the lowest-ratio efficiency (E_{22}, the smallest) are varied and opposite to each other. The gap between the two varies significantly in the range of 0.057–0.559. This observation reveals the fact that the dynamics of the environmental performance may be viewed quite differently depending on the objectives of the analysis.

The weighted and average efficiencies provide a more objective picture of the environmental performance in dynamics. In our study, they were pretty close to each other and had an increasing trend beyond 2000.

Figure 4. Environmental and economic ratios in the dynamics of energy consumption and energy-related CO_2 emissions. Graphs of ratios for 1980–2010 calculated from Table 4 for GDP over energy (upper left), GDP over population (upper right), translated CO_2 emissions over energy (lower left), and translated CO_2 emissions over population (lower right).

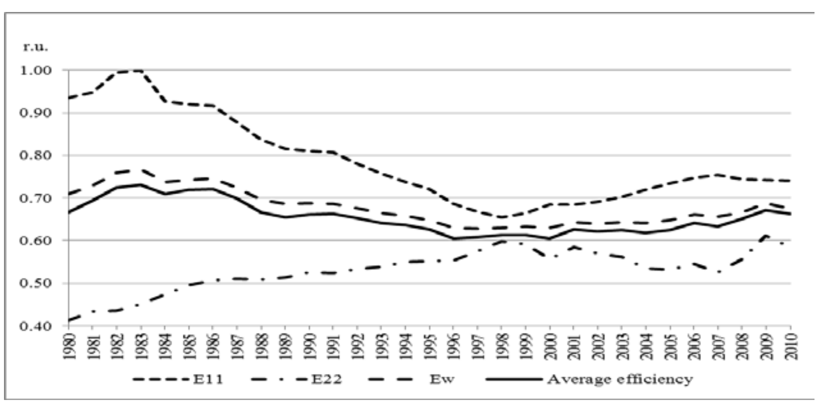

Figure 5. Efficiencies in the dynamics of energy consumption and energy-related CO_2 emissions. Graphs of results for 1980-2010 from Table 5 for DEA PO (E_{11}), Lowest - Ratio (E_{22}), Weighted (E_w), and the average.

Quantitative Reasoning

We can continue with other examples from different subject areas to demonstrate that the variety of possible applications of DEA and DEA PO is practically unlimited. The numbers of inputs and outputs are arbitrary; we used two inputs and two outputs in this paper merely for ease of calculation. The following are some possible applications, just to name a few:

- Company valuation, with profits, sales, and market capitalization as outputs, and equity and debt as inputs.
- College comparisons, with the number of students and graduation rate as outputs, and the amount of faculty, staff, operational expenses, and tuition fees as inputs.
- Car appraisal, with miles per gallon and safety scores as output, and price and maintenance expenses as inputs.

Studying DEA with such applications has the potential to provide students practice in critical thinking and quantitative reasoning in authentic contexts. It can provide them the opportunity to learn to determine positive and negative factors (outputs and inputs, respectively) and to weigh the two groups against each other. Moreover, DEA PO processes objects independently, one at a time, thus allowing potentially for subdividing large tasks among working groups of students who come together after their work to combine their results for analysis and discussion.

Conclusions

From the two examples examined in this paper, Data Envelopment Analysis with a Perfect Object (DEA PO) appears to have promise as a teaching and analytical tool. DEA PO adds a virtual Perfect Object—one with the largest outputs and smallest inputs—to a collection of actual DEA objects subject to the evaluation of their efficiency of operation. Including the PO allows one to use simple, explicit formulas for efficiency calculations. Since DEA PO evaluates efficiency based on the best output-to-input ratio only, it is extended further in this paper to DEA PO with Partial Perfect Objects (i.e., to DEA PPO), by which procedure a collection of PO's is generated by the sequential elimination of larger outputs and smaller inputs. The resulting collection of the PPO's exhausts all possible output-to-input ratios, so that the corresponding partial efficiencies provide a wider base on which to evaluate the efficiency and rank the objects.

Applying the DEA PO and PPO to the 2012 Olympic Games in London, we have demonstrated a way to find the winningest country in the Olympic Games

and rank the national Olympic teams. The investigation into the dynamics of the United States' environmental efficiency revealed how well the CO_2 emissions were adjusted for GDP, energy consumption, and population in the period of 1980–2010.

The suggested approach may be used in different situations where several types of resources are involved resulting in a series of different outcomes. The multitude of potential applications combined with the simplicity of the computations makes this approach attractive for educational purposes, in particular for Quantitative Reasoning courses at the undergraduate college level and for enrichment courses in high school.

Acknowledgments

This research was partially supported by a PSC-CUNY Award # 65135-00 43, jointly funded by The Professional Staff Congress and The City University of New York. The Author is grateful to the three anonymous reviewers for profound analysis of the initial version of the paper and fruitful comments; special thanks to Reviewer 3 and to the Editor

References

Andersen, P., and N. Petersen. 1993. A procedure for ranking efficient units in Data Envelopment Analysis. *Management Science* 39 (10): 1261–1264. http://dx.doi.org/10.1287/mnsc.39.10.1261

Banker, R., A. Charnes, and W. Cooper. 1984. Some models for estimating technical and scale efficiency in Data Envelopment Analysis. *Management Science* 30 (9): 1078–1092. http://dx.doi.org/10.1287/mnsc.30.9.1078

Charnes, A., W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision-making units. *European Journal of Operational Research* 2: 429–444. http://dx.doi.org/10.1016/0377-2217(78)90138-8

Cook W. and J. Zhu. 2006. Incorporating multi-process performance standards into the DEA Framework. *Operations Research* 54(4): 656–665. http://dx.doi.org/10.1287/opre.1060.0295

Cooper W., L. Seiford, and J. Zhu (eds.). 2011. *Handbook on data envelopment analysis*. International Series in Operations Research & Management Science, vol 164, 2nd ed. New York: Springer.

Daneshian, B., G. Jahanshahloo, F. Lotfi, T. Allahviranloo, and S. Mehrabian. 2005. The overall assurance interval for the non-Archimedean epsilon in DEA models. *Mathematical and Computational Applications* 10(3): 387–393.

Färe, R., S. Grosskopf, S., C. Lovell, and C. Pasurka. 1989. Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach. *The Review of Economics and Statistics* 71: 90–98. http://dx.doi.org/10.2307/1928055

Farrell, M. 1957. The measurement of production efficiency. *Journal of the Royal Statistical Society*, Series A, 120 (3): 253–282. http://dx.doi.org/10.2307/2343100

Golany, B. 1988. An interactive MOLP procedure for the extension of DEA to effectiveness analysis. *Journal of Operational Research Society* 39: 725–734. http://dx.doi.org/10.2307/2583767

Seiford L., and Zhu J. 2002. Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research* 142: 16–20. http://dx.doi.org/10.1016/S0377-2217(01)00293-4

Thanassoulis E., and R. Dyson. 1992. Estimating preferred target input-output levels using Data Envelopment Analysis. *European Journal of Operational Research* 56: 80–97. http://dx.doi.org/10.1016/0377-2217(92)90294-J

Vaninsky, A. 2009. Environmental erformance 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*) 54:139–144. http://www.waset.org/journals/waset/v30/v30-26.pdf (accessed April 8, 2013) .

———. 2011a. DEA with a perfect object: Analytical Solutions. *Communications in Mathematics and Applications* 2(1): 1–13.

———. 2011b. Explicit formulas for efficiency scores and weight coefficients in DEA problems with a perfect object. *International Journal of Mathematical Modeling, Simulation and Applications* 4(3): 217–233.

Zhou P., B. Ang, and K. Poh. 2008. Measuring environmental performance under different environmental DEA technologies. *Energy Economics* 30: 1–14. http://dx.doi.org/10.1016/j.eneco.2006.05.001

Zhu, J. 1996. Data Envelopment Analysis with preference structure. *Journal of the Operational Research Society* 47: 136–150. http://dx.doi.org/10.2307/2584258

Appendix. Data Envelopment Analysis with a Perfect Object

In this section, we describe Data Envelopment Analysis (DEA) and introduce DEA with a Perfect Object (DEA PO), following Vaninsky (2011a). DEA was developed in Charnes et al. (1978) and Banker et al. (1984). It has become a well-established tool to estimate the relative efficiency of a group of objects referred to

as Decision-Making Units (DMUs) that use inputs $\mathbf{X} = (X_j, j = 1,...,r) > 0$ to produce outputs $\mathbf{Y} = (Y_i, i = 1,..., s) > 0$. DEA combines all of the indicators of each object into a single efficiency score scaled to an interval [0, 1]. An object is considered efficient if it receives a score equal to 1 and inefficient if it receives a score of less than 1. The DEA efficiency score is based on the efficiency ratio suggested in Farrell (1957):

$$E = \frac{\sum_{i=1}^{s} u_i Y_i}{\sum_{j=1}^{r} v_j X_j}, \qquad \text{(A1)}$$

where $\boldsymbol{u} = (u_1,..., u_s) \geq 0$ and $\boldsymbol{v} = (v_1,..., v_r) \geq 0$ are nonnegative weights assigned to outputs and inputs, respectively. The solution to the problem (A1) is defined up to proportional change in the weights $\boldsymbol{u}$ and $\boldsymbol{v}$.

DEA begins with a problem that is seemingly different from that given by (A1). For each DMU$_k$ taken in turn, it builds a virtual DMU as a linear combination of all DMU's in the group that uses at most $\theta_k \leq 1$ share of each input, while producing at least the same amount of outputs. Mathematically, DEA sets up a series of the Linear Programming (LP) optimization problems:

For each DMU$_k$, $k = 1,\ldots, N$, find a nonnegative vector $\boldsymbol{\lambda}_k=(\lambda_{k0}, \lambda_{k1}, \lambda_{k2},\ldots, \lambda_{kN}) \geq 0$ and scalar θ_k that

minimize θ_k,
subject to

$$\sum_{m=0}^{N} \lambda_{km} X_{mj} \leq \theta_k X_{kj},\ j = 1,...,r;$$
$$\sum_{m=1}^{N} \lambda_{km} Y_{mi} \geq Y_{ki},\ i = 1,...,s; \qquad \text{(A2)}$$
$$\lambda_{km} \geq 0,\ m = 1,...,N.$$

This is the main DEA input minimization model with constant returns to scale (IM CRS). The minimum value of θ_k is called efficiency of the DMU$_k$.

DEA proceeds with an LP problem dual to (A2) that, after elementary algebraic transformations, is this:

For each DMU$_k$, $k = 1,\ldots,N$

$$maximize \sum_{i=1}^{s} u_{ki} Y_{ki}$$

subject to

$$\sum_{i=1}^{s} u_{ki} Y_{mi} \le \sum_{j=1}^{r} v_{kj} X_{mj}, \; m = 1,\ldots, N \qquad \text{(A3)}$$

$$\sum_{j=1}^{r} v_{kj} X_{kj} \le 1,$$

see Wikipedia[13] for detail.

It was shown in Charnes et al. (1978) that the dual optimization problem (A3) is equivalent to the problem (A1) with u_{ki} and v_{kj} playing the roles of the weight coefficients, appended with an additional condition normalizing the weight coefficients v_{kj}:

$$\sum_{j=1}^{r} v_{kj} X_{kj} = 1, \qquad \text{(A4)}$$

This finding allowed for the rewriting of model (A1) as follows:

For each DMU$_k$, $k = 1,\ldots, N$, find nonnegative vectors $\boldsymbol{u}_k = (u_{k1},\ldots, u_{ks}) \ge 0$ and $\boldsymbol{v}_k = (v_{k1},\ldots, v_{kr}) \ge 0$ to

$$maximize\ E_k = \frac{\sum_{i=1}^{s} u_{ki} Y_{ki}}{\sum_{j=1}^{r} v_{kj} X_{kj}}, \qquad \text{(A5)}$$

subject to

$$\sum_{j=1}^{r} v_{kj} X_{kj} = 1, \qquad \text{(A6)}$$

and $E_m \le 1$ for all DMU$_m$, $m = 1,\ldots, N$ taken with the same weight coefficients $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$. Restriction (A6) imposed on the weights $\boldsymbol{v}$ provides the uniqueness of the solution. The optimal value of the LP problem is equal to that of its dual problem, $\theta_k = E_k$.

---

[13] "Linear programming." Wikipedia. http://en.wikipedia.org/wiki/Linear_programming (accessed April 5, 2013).

Geometrically, the model (A2) measures efficiency as a relative *horizontal* distance to the frontier of possibilities. There is another approach to measuring the efficiency as a ratio of relative *vertical* distances to the point and to the frontier. It results in the output-maximization DEA model with constant returns to scale - OM CRS. It is known that for the CRS models the efficiency scores are equal: $E_{IM} = E_{OM}$, and do not depend on the units of measurement of inputs or outputs. Some of other DMU models do not hold these properties.

DEA is a non-parametric method; it does not require an *apriori* functional relationship among the inputs and outputs. One of the main advantages of DEA in its ratio form (A5) is its capacity to objectively assign values to the weight coefficients $\boldsymbol{u}$ and $\boldsymbol{v}$. Conceptually, DEA allows each DMU to assign the weight coefficients to each input and output favorably. However, the potential of a given DMU to achieve the maximum efficiency score is bound by the requirement that with the weight coefficients assigned to it, no other DMU in the group receives an efficiency score greater than 1. This means that a poorly performing DMU cannot achieve a high efficiency score for itself through a manipulation of the weight coefficients. If this were the case, an object that performs really well would have received an efficiency score greater than 1.

DEA PO appends conventional DEA with a virtual Perfect Object (PO) - the one having the largest outputs and smallest inputs. The PO serves as a universal benchmark for the DMU's comparisons. In the framework of DEA PO, all inputs and outputs are transformed to their ratios to the corresponding PO indicators. Fig. (6), similar to that in Vaninsky (2009), gives a geometric interpretation. In the figure, the Perfect Object is located at the point F so that $X_{0j} = min\ X_{kj}$, $j = 1,\ldots, r$; $Y_{0i} = max\ Y_{ki}$, $i = 1,\ldots, s$; $k = 1, \ldots, N$. DEA frontier passes through the point B, the location of the DMU with a maximum ratio of output to input. DEA PO frontier passes through the point F corresponding to the maximum output and minimum input in the group. For a DMU_k located at the point G, the DEA efficiency score equals to HL/HG, while the DEA PO efficiency score is HK/HG.

As was shown in Vaninsky (2011a), in the presence of the Perfect Object, the efficiency score may be found as a ratio of the largest relative output to the smallest relative input:

$$E_k = \max_{0\le j\le r} \frac{X_{0j}}{X_j} \times \max_{0\le i\le s} \frac{Y_i}{Y_{oi}} = \frac{X_{0j*}Y_{i*}}{X_{j*}Y_{0i*}} = \frac{\left(Y_{i*} \middle/ Y_{0i*} \right)}{\left(X_{j*} \middle/ X_{0j*} \right)}$$

$$= \frac{\textit{maximum relative output}}{\textit{minimum relative input}} \qquad \text{(A7)}$$

where lower indexes i = 1..s and j = 1..r stand for outputs and inputs, respectively, lower indexes i^* and j^*, for the output and input providing the maximum value of the output and the minimum value of the input ratios, correspondingly; lower index *0* stands for the Perfect Object. The last term of the Equation A7 is used as Equation 1 in the text.

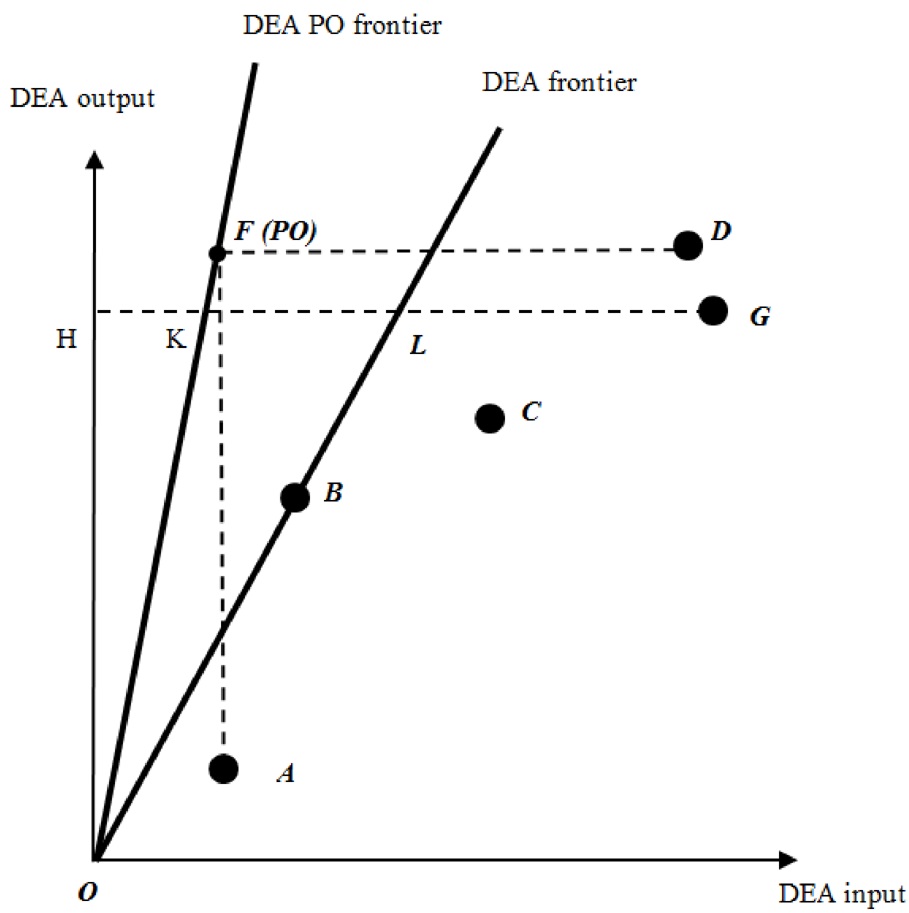

Figure 6. DEA frontiers for one input and one output.
Points *A, B, C, D*, and *G* are the locations of the DMU's. *OB* is a DEA CRS frontier, *OF*, is a DEA PO frontier. A Perfect Object is located at point *F*, corresponding to minimum input and maximum output in the group. All actual objects are located to the right or on the DEA frontier. The DEA frontier passes through the point of location of the DMU with a maximum ratio of output to input. The DEA PO frontier passes through the point of maximum output and minimum input in the group. Input-oriented DEA efficiency score at point *G* equals to *HL/HG,* DEA PO efficiency score, *HK/HG.*

This version of the DEA PO corresponds to the DEA CRS models and combines both IM and OM ones. In particular, the DEA PO eliminates the need in the ranking of efficient objects, Andersen & Petersen (1993).

As follows from the Equations A2 and A3, conventional DEA, methodologically, assigns efficiency scores based on smaller inputs and larger outputs. In particular, an object having only one smallest input or one largest output in the group will receive an efficiency score of 1, and thus, will be considered efficient no matter what all of its other inputs or outputs are. In the dual LP problem (A3), only those weight coefficients u_{ki} and v_{kj} that correspond to the equalities in (A2) at the optimum point are positive. All other weight coefficients are equal to zero. This means that a DMU$_k$ can arbitrarily increase the inputs or decrease outputs related to zero weight coefficients, while keeping its efficiency score unchanged. DEA suggests different modifications to the basic models (A2) and (A3) to avoid these undesirable properties, if this property is not acceptable.

DEA PO inherits the DEA methodology and even hardens it by using the ratio of the best relative output to the best relative input only. In case that such an approach may create a problem, DEA PO is able to address it by using *partial PO*'s, *partial efficiencies,* and *weighted efficiency,* as suggested in this paper. The partial efficiency E_{kl} is the efficiency of a Partial Perfect Object (PPO) obtained by eliminating (k–1) largest outputs and (l–1) smallest inputs, as given by the Equation 2 in the text. The weighted efficiency is a weighted arithmetic average of the partial efficiencies:

$$\sum_{k=1}^{s}\sum_{l=1}^{r} w_{kl} E_{kl} \,, \tag{A8}$$

where

$$\sum_{k=1}^{s}\sum_{l=1}^{r} w_{kl} = 1 \,. \tag{A9}$$

This paper suggests assigning the weight coefficients w_{kl} proportional to the share of the total number of inputs and outputs that the corresponding kl-partial perfect object comprises:

$$w_{kl} = W \times \frac{(s-k-1)+(r-l-1)}{r+s} = W \times \left(1 - \frac{k+l-2}{r+s}\right) \,, \; k = 1..s, \; l = 1..r, \tag{A10}$$

where r and s stand for the total number of inputs or outputs, respectively, W, for a normalizing coefficient providing the sum of the weights equals to 1:

$$W = \frac{1}{\sum_{k=1}^{s}\sum_{l=1}^{r}\left(1-\frac{k+l-2}{r+s}\right)} \quad . \tag{A11}$$

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

$$\begin{gathered}\sum_{k=1}^{s}\sum_{l=1}^{r}\left(1-\frac{k+l-2}{r+s}\right) = \frac{1}{r+s}\times\sum_{k=1}^{s}\sum_{l=1}^{r}(r+s+2-k-l) = \\ \frac{1}{r+s}\times\left(\sum_{k=1}^{s}\sum_{l=1}^{r}(r+s+2)-\sum_{k=1}^{s}\sum_{l=1}^{r}r-\sum_{k=1}^{s}\sum_{l=1}^{r}s\right) = \\ \frac{rs(r+s-2)}{2(r+s)} \quad ,\end{gathered} \tag{A12}$$

so that

$$W = \frac{2(r+s)}{rs(r+s-2)}. \tag{A13}$$

The computations in Equation A12 leading to the Equation A13 may be easily carried out by using a Computer Algebra System, such as that of the graphing calculator TI-89. Substituting (A13) into (A10), we get:

$$\begin{gathered}w_{kl} = W\times\frac{(s-k-1)+(r-l-1)}{r+s} = \frac{2(r+s)}{rs(r+s-2)}\times\frac{(s-k-1)+(r-l-1)}{r+s} = \\ \frac{2(r+s-k-l+2)}{rs(r+s+2)}.\end{gathered} \tag{A14}$$

The weighted efficiency scores w_{kl} in Tables 3 and 5 were obtained by using Equation A14 with $r = s = 2$.