Omega 44 (2014) 1–4

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

Technical note

Data envelopment analysis: Prior to choosing a model

Wade D. Cook [a], Kaoru Tone [b], Joe Zhu [c,*]

[a] Schulich School of Business, York University, Toronto, Ontario, Canada M3J 1P3
[b] National Graduate Institute for Policy Studies, Minato-ku, Tokyo 106-8677, Japan
[c] School of Business, Worcester Polytechnic Institute, Worcester, MA 01609, USA

ARTICLE INFO

Article history:
Received 28 April 2013
Accepted 26 September 2013
Available online 7 October 2013

Keywords:
Data envelopment analysis (DEA)
Efficiency
Input
Output
Ratio

ABSTRACT

In this paper, we address several issues related to the use of data envelopment analysis (DEA). These issues include model orientation, input and output selection/definition, the use of mixed and raw data, and the number of inputs and outputs to use versus the number of decision making units (DMUs). We believe that within the DEA community, researchers, practitioners, and reviewers may have concerns and, in many cases, incorrect views about these issues. Some of the concerns stem from what is perceived as being the purpose of the DEA exercise. While the DEA frontier can rightly be viewed as a production frontier, it must be remembered that ultimately DEA is a method for performance evaluation and benchmarking against best-practice. DEA can be viewed as a tool for multiple-criteria evaluation problems where DMUs are alternatives and each DMU is represented by its performance in multiple criteria which are coined/classified as DEA inputs and outputs. The purpose of this paper is to offer some clarification and direction on these matters.

1. Introduction

Data envelopment analysis (DEA) is a mathematical programming based approach for measuring relative efficiency of decision making units (DMUs) that have multiple inputs and outputs [7]. Whether it is the researcher, the practitioner or the student, the use of the DEA methodology gives rise to some important questions before proceeding to a DEA analysis:

"What is the purpose of the performance measurement and analysis?"
"What are the decision-making units (DMUs) and the outputs and inputs to be used to characterize the performance of those DMUs?"
"What is an appropriate number of DMUs, given the number of inputs and outputs chosen?"
"What is the appropriate model orientation (input, output, additive)?"
"Does the analysis involve the use of ratio and raw data in the same model, and is this appropriate?"

We believe that within the DEA community, researchers, practitioners, and reviewers may have concerns and often incorrect views about these issues. The evidence for this concern materializes primarily in unpublished (and confidential) referee reports. It is, for example, very common practice, particularly in the case of novice users, to invoke the input-oriented constant returns to scale model in cases where inputs are in fact not under management control. More will be said regarding this below. Some of the concerns stem as well from what is perceived as being the purpose of the DEA exercise. While the DEA frontier can, in some situations, be rightly viewed as a production frontier, it must be remembered that ultimately DEA is intended as a method for performance evaluation and benchmarking against best-practice. The purpose of this paper is to offer some clarification and direction on some of these matters.

In the sections to follow we attempt to provide some guidance on, and possibly some answers to these questions.

2. Purpose of the performance measurement exercise

In any study of organizational efficiency it is necessary to have a clear understanding of the "process" being evaluated. A study of hospital efficiency, for example, must provide clarity as to which elements of the organization are being evaluated. Is it particular wards in the chosen hospitals (e.g. maternity wards), or particular functions such as emergency room procedures, or is it the cost effectiveness of the entire organization that is at issue? A clear specification of the function to be studied will drive the choice of inputs and outputs to be examined. A recent study of schools of business by Aviles-Sacoto [2], for example, placed considerable emphasis on the data gathering exercise aimed at understanding

* Corresponding author. Tel.: +1 508 831 5467.
E-mail addresses: wcook@schulich.yorku.ca (W.D. Cook), tone@grips.ac.jp (K. Tone), jzhu@wpi.edu (J. Zhu).

0305-0483/$ - see front matter
http://dx.doi.org/10.1016/j.omega.2013.09.004

the precise measures deemed important by management. In that specific case it was international internships and job success, on the part of students, that were two of the most important factors for capturing school reputation. As discussed below, the "purpose" of the performance measurement exercise influences the model orientation. Numerous such examples are in abundance in the literature. (See, e.g., [3,9,19,27].)

3. DEA inputs and outputs

In the literature, DEA is generally introduced as a mathematical programming approach for measuring relative efficiencies of DMUs, when multiple inputs and multiple outputs are present. While the concept of inputs and outputs is well understood, it is often the case that researchers take the notion for granted, and little attention tends to be paid to insuring that the selected measures properly reflect, to the greatest extent possible, the "process" under study. While it is the case, as with regression analysis, that one can never be completely assured that all of the relevant variables have been included, every attempt should be made to include those that make practical sense for the setting under investigation. As a case in point, the original DEA model of Charnes, Cooper and Rhodes [7,8], involving the study of school districts in Texas, was developed in a ratio form of outputs/inputs, but the authors provide little in the way of rationalization in regard to appropriate variables (inputs and outputs) for studying student performance. This is not to imply that the variables used were not appropriate for the problem at hand, but rather it serves to illustrate that the paper, like many of those that followed over the past three decades, was primarily focused on methodological development. One gets the sense in much of the literature that there is little need to spend time laboring over how a process actually works. After all, in a production or service process, inputs and outputs are generally clearly defined. For example, the number of employees and profits are obvious examples of an input and an output, respectively.

Although DEA has a strong link to production theory in economics, the tool is also used for benchmarking in operations management, where a set of measures is selected to benchmark the performance of manufacturing and service operations. In the circumstance of benchmarking, the efficient DMUs, as defined by DEA, may not necessarily form a "production frontier", but rather lead to a "best-practice frontier". For example, if one benchmarks the performance of computers, it is natural to consider different features (screen size and resolution, memory size, process speed, hard disk size, and others). One would then have to classify these features into "inputs" and "outputs" in order to apply a proper DEA analysis. However, these features may not actually represent inputs and outputs at all, in the standard notion of production. In fact, if one examines the benchmarking literature, other terms, such as "indicators", "outcomes", and "metrics", are used. The issue now becomes one of how to classify these performance measures into inputs and outputs, for use in DEA.

In general, DEA minimizes "inputs" and maximizes "outputs"; in other words, smaller levels of the former and larger levels of the latter represent better performance or efficiency. This can then be a rule for classifying factors under these two headings. There are, however, exceptions to this; for example, pollutants from a production process are outputs, yet higher levels of these indicate worse performance. There are DEA models that deal with such *undesirable* outputs (see, e.g., [21,17].)

In certain circumstances, a factor can play a dual role of input and output simultaneously. For example, when evaluating the efficiencies of a set of universities, if one considers the numbers of Ph.D. students trained as outcomes from the education process, then this factor can rightly be viewed as an output. At the same time, however, Ph.D. students assist in carrying out research, and can therefore be viewed as a resource, hence an input to the process. See [12]. In such cases, the user must clearly define the purpose of benchmarking so that such performance measures can be classified as inputs or outputs. In some situations, the DMUs may have internal structures, e.g., a two-stage process. For example, banks generate deposits as an output in the first stage, and then the deposits are used as an input to generate profit in the second stage. In this case, "deposits" is treated as both output (from the first stage) and input (to the second stage).

In summary, if the underlying DEA problem represents a form of "production process", then "inputs" and "outputs" can often be more clearly identified. The resources used or required are usually the inputs and the outcomes are the outputs. If, however, the DEA problem is a general benchmarking problem, then the inputs are usually the "less-the-better" type of performance measures and the outputs are usually the "more-the-better" type of performance measures. The latter case is particularly relevant to the situations where DEA is employed as a MCDM (multiple criteria decision making) tool (see, e.g., [5,14,24]). DEA then can be viewed as a multiple-criteria evaluation methodology where DMUs are alternatives, and DEA inputs and outputs are two sets of performance criteria where one set (inputs) is to be minimized and the other (outputs) is to be maximized. In DEA, these multiple criteria are generally modeled as in a ratio form, e.g., the CCR ratio model [7]

$$\max e_{j_o}$$

$$\text{subject to } e_j < 1 \tag{1}$$

where

$$e_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}$$

and x_{ij} and y_{rj} represents DEA *inputs and outputs*, and v_i and u_r are unknown weights. Obviously, x_{ij} and y_{rj} can be referred to in different terms, rather than "inputs" and "outputs". Inputs may, for example, be quality measures that act as surrogates for resources expended by the DMU. Outputs may, as well, appear in the form of outcomes such as employee satisfaction.

4. The numbers of inputs and outputs

It is well known that large numbers of inputs and outputs compared to the number of DMUs may diminish the discriminatory power of DEA. A suggested "rule of thumb" is that the number of DMUs be at least twice the number of inputs and outputs combined (see [16]). Banker et al. [4] on the other hand state that the number of DMUs should be at least three times the number of inputs and outputs combined. However, such a rule is neither imperative, nor does it have a statistical basis, but rather is often imposed for convenience. Otherwise, it is true that one loses discrimination power. It is not suggested, however, that such a rule is one that must be satisfied. There are situations where a significant number of DMUs are in fact efficient. In some cases the population size is small and does not permit one to add actual DMUs beyond a certain point. However, if the user wishes to reduce the number or proportion of efficient DMUs, various DEA models can help; for example, weight restrictions may be useful in such cases.

We point out that while in statistical regression analysis, sample size can be a critical issue, as it tries to estimate the average behavior of a set of DMUs, DEA when used as a benchmarking tool, focuses on individual DMU performance. In that sense, the size of the sample or the number of DMUs under evaluation may be immaterial.

For example, if there are only 10 firms in a particular market and if a large number of inputs and outputs have to be used if deemed necessary by the management, then the benchmarking results obtained from DEA can still be of value. One fact remains, namely that whatever form the production frontier takes, it is beyond the best practice frontier. It is also true that if one adds an additional DMU to an existing set, that DMU will be either inefficient or efficient. In the former case, the best practice frontier does not shift, and nothing new is learned about the production frontier. In the latter situation, the frontier may shift closer to the actual (but unknown) production frontier.

In summary, DEA is not a form of regression model, but rather it is a frontier-based linear programming-based optimization technique. It is meaningless to apply a sample size requirement to DEA, which should be viewed as a benchmarking tool focusing on individual performance. It is likely that a significant portion of DMUs will be deemed as efficient, if there are "too many" inputs and outputs given the number of DMUs. If the goal is to obtain fewer efficient DMUs, then one can use weight restrictions or other DEA approaches to reduce the number of efficient DMUs (see, e.g., [1,11,13,25]).

Finally, it is important to capturing as many as possible of the relevant inputs and outputs for a DEA analysis. While there is no magic formula or model that can guarantee one has all the performance measures, Golany and Roll [16] provide a detailed procedure on the selection of DEA inputs and outputs and DMUs. Users may also find that empirical survey papers help in identifying DEA inputs and outputs. For example, Avkiran [3], Chilingerian and Sherman [9], Paradi and Zhu [19], and Triantis [27] provide detailed documentations on DEA empirical applications in banking, health care, engineering, and service sectors in general.

5. Orientation

A DEA analysis should clearly identify what is to be achieved from that analysis. Consider an efficiency study of hospitals where inputs are such factors as numbers of bed days available and hospital budget, while outputs are numbers of patients served and numbers of nurses trained. If the goal is to identify units that are over-utilizing resources, then it would appear that input reduction is to be the central focus of the exercise. In such a situation, the input-oriented DEA model would seem to be appropriate. On the other hand, in the business school study mentioned above, where input are quality measures earned by the school and the students (percent of students entering with scholarships, and the school's academic rating), and outputs are achievements in the form of internships and jobs, one can argue that it is output enhancement, not input reduction on which management will focus attention. This would appear to point to the output-oriented model as the appropriate analysis tool. One might even question the orientation logic of the Charnes et al. [8] paper. The use of the input oriented model in that case could be questionable. It would seem to imply that input reduction (where inputs were such things as average supervision hours supplied by the parents of the students) is the logical course of action that an inefficient school district should take to become efficient. It would seem more appropriate to call upon school districts to enhance grade performance measures (the outputs), hence the output-oriented model would appear to make more sense.

If both input reduction and output enhancement are desirable goals in a particular application, then a slacks-based measure [26] may provide the appropriate model structure to capture a DMU's performance measure.

Thus, the analyst needs to articulate the purpose of the analysis, whether input reduction, output expansion or both. However, we should note that from a DEA modeling point of view, both input and output orientations will yield the same efficient or best practice frontier under a specific returns to sale (RTS) assumption, for example, constant or variable returns to scale. Therefore, if it is the best practice that is of interest, orientation does not matter. However, in this case, the efficient reference sets for inefficient units may differ.

Finally, the orientation of a model is not so obvious to the casual user, if one simply observes the ratio forms of DEA models, e.g., the CCR ratio or CRS model (1). It is not immediately clear why one would refer to model (1) an input-oriented model. To solve (1), one normally transforms it into the following linear program using the Charnes–Cooper transformation [6],

$$\max \sum_{r=1}^{s} \mu_r y_{rj_o}$$

subject to

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \omega_i x_{ij} \le 0 \qquad (2)$$

$$\sum_{i=1}^{m} \omega_i x_{ij_o} = 1$$

The dual form of model (2) is then given by:

$$\min \theta$$

subject to

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{ij_o} && i = 1, \ldots, \text{m} \\ &\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} && r = 1, \ldots, \text{s} \\ &\lambda_j \ge 0 && j = 1, \ldots, \text{n} \end{aligned} \qquad (3)$$

Note that it is also not obvious that model (2) is input-oriented, since a novice DEA user may view it rather as being output-oriented, given that the objective function appears in the form of aggregated outputs. It is only when one turns to the dual of (2), namely (3), with the compression of inputs by the factor θ, that the reference to input orientation is made more transparent.

Note also that an output-oriented version of model (1) is to minimize the inverse ratio $(1/e_{j_o})$. However, such a transformation does not obviously imply a change of orientation.

6. Mixed use of ratio and raw data

Dyson et al. [15] discussed several pitfalls in DEA. One of the pitfalls is that the efficiency score can be misjudged, when input and output variables simultaneously take the form of percentiles and/or ratios (e.g., profit per employee, and returns on investment), and raw data (e.g., revenues, assets, employees, profits). [10] provide an example to demonstrate the potential problem when ratio and raw data are mixed in a DEA analysis.

However, we would like to point out that a mixture of ratios/percentiles and raw data is permissible in DEA applications. It is too restrictive to reach the conclusion that these two forms of data cannot coexist in a model. To demonstrate, let us first revisit the example provided in [10]. This example asks the reader to criticize the selection of two inputs, the annual average salary per employee and the number of employees, and one output given by the annual average sales per employee. As illustrated in [10], for DMU_i, let p_i be the number of employees, c_i be the total annual salary, and d_i be the total sales. Then the DEA ratio can be expressed as $u_1(d_i/p_i)/(v_1(c_i/p_i)+v_2p_i)$ where u_1, v_1 and v_2 are weights. This ratio is equivalent to $u_1d_i/(v_1c_i+v_2p_i^2)$. Therefore, this particular example demonstrates that emphasis is put on the number of employees by squaring this value, while other factors are evaluated in a linear manner.

This example indicates that when a factor (e.g., number of employees) appears both at the input and output sides, one may have an issue, perhaps only when ratio and raw data are not properly mixed. The example is not intended to show that one cannot mix ratio and raw data. In fact, Cooper, Seiford and Tone [10] further points out that that if such uneven treatment has no special justification, it may be better to use total sales (d_i) as output and total salary (c_i) and number of employees (p_i) as two inputs. There may of course be situations where such justification can be provided. For instance, in this particular example, we can redefine the weight as $\hat{v}_2 = (v_2 \times p_i)$. As a result, we do not have a squared term of p_i. The newly defined weight $\hat{v}_2$ will be different for each DMU, reflecting a user's preference or weight restrictions. As a result, we obtain a different DEA model.

While Cooper, Seiford and Tone [10] state that one should be careful in dealing with process data and raw data at the same time, their example does not appear to provide any solid justification for not using these two forms of data simultaneously. It is not clear that one can generalize this one specific example to cases involving a mixture of ratio and raw data.

We next point out that a problem with index measures may arise in CRS as discussed in [15]. If the index value is the same across all DMUs, or more generally, if one input (in an output-oriented CRS model) or one output (in an input-oriented CRS model) has an equal value across all DMUs, the CRS becomes a variable returns to scale (VRS) model. This is because the related input or output constraint becomes the convexity constraint in the CRS model (see, e.g., Theorem 3 in [20]).

We finally should point out that in the VRS model, if the ratio data are in percentages, the DEA projections remain in the range (0%, 100%). However, in the constant returns to scale (CRS) setting, the same situation is not always true. In particular, in an output-oriented CRS model, the projection of a percentage output can go above 100%. The user should, therefore, exercise caution when using a CRS technology.

7. Conclusions

Despite the many applications of DEA that have been advanced in the literature, it would appear that in many cases little attention is paid to a number of important modeling issues. Some of these pertain to clearly specifying the purpose of the analysis, deciding on inputs and outputs, choosing a model orientation, and paying attention to the type of data involved, whether ratio versus raw data. The primary purpose of this paper is to raise awareness of these issues, and to offer some advice and opinions in that regard. As pointed in a citation-based DEA survey by Liu et al. [18], it is expected that the literature will grow to at least double its current size. Therefore, it is critical that the DEA community has an open mind on these issues, as DEA is being further developed and applied in various areas.

We would like to point out that if DEA is used as a pure benchmarking approach (without having a real production function), the meaning of efficiency as a distance to the frontier may no longer be valid. However, DEA still yields information on relative distance to the best-practices. We note that a user needs to exercise caution if he/she uses DEA-based returns-to-scale (RTS) identification, scale efficiency, Malmquist productivity index, and others where the DEA model is intended to have a production meaning as a prerequisite. Also, under general benchmarking, the DEA score may no longer be referred to as "production efficiency". In this case, we may wish to refer to the DEA score as a form of "overall performance" of an organization. Such "overall performance" can appear in the form of composite measure that aggregates individual indicators (inputs and outputs) via a DEA model. For example, composite measures (DEA scores) of quality indicators allow senior leaders to better benchmark their organization's performance against other high-performing organizations [23].

In conclusion, we believe that DEA inputs and outputs can represent more than the usual concept of "inputs" and "outputs" in a conventional production process; and DEA is more than an efficiency measure under the notion of a production process. In addition to being used as an estimate of production efficiency, DEA is a type of "balanced benchmarking" [22] that examines performance in multiple criteria and helps organizations to test their assumptions about performance, productivity, and efficiency.

References

[1] Allen R, Athanassopoulos A, Dyson RG, Thanassoulis E. Weights restrictions and value judgments in data envelopment analysis: evolution, development and future directions. Annals of Operations Research 1997;73:13–34.
[2] Aviles-Sacoto S. A restricted multiplier DEA model for identifying best practices for attracting students into a Mexican university- case study: School of Business at ITESM, Master's thesis, (ITESM) (Technological and Higher Education Institute of Monterrey, Mexico); 2012.
[3] Avkiran NK. Applications of data envelopment analysis in the service sector. In: Cooper WW, Seiford LM, Zhu J, editors. Handbook on data envelopment analysis. New York: Springer; 2011. p. 2011.
[4] Banker RD, Charnes A, Cooper WW, Swarts J, Thomas D. An introduction to data envelopment analysis with some of its models and their uses. Research in Government and Nonprofit Accounting 1989;5:125–163.
[5] Belton V, Vickers SP. Demystifying DEA: a visual interactive approach based on multi criteria analysis. Journal of the Operational Research Society 1993;44 (9):883–896.
[6] Charnes A, Cooper WW. Programming with linear fractional functionals. Naval Research Logistics Quarterly 1962;9:181–185.
[7] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research 1978;2:429–444.
[8] Charnes A, Cooper WW, Rhodes E. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. Management Science 1981;27(6):668–697.
[9] Chilingerian JA, Sherman HD. Health-care applications: from hospitals to physicians, from productive efficiency to quality frontiers. In: Cooper WW, Seiford LM, Zhu J, editors. Handbook on data envelopment analysis. New York: Springer; 2011. p. 2011.
[10] Cooper WW, Seiford LM, Tone K. Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver Software, second edition; 2007. p19.
[11] Cooper WW, Ruiz JL, Sirvent I. Choice and uses of DEA weights. In: Cooper WW, Seiford LM, Zhu J, editors. Handbook on data envelopment analysis. New York: Springer; 2011. p. 2011.
[12] Cook WD, Green R, Zhu J. Dual-role factors in DEA. IIE Transactions 2006;38:1–11.
[13] Cook WD, Zhu J. Context-dependent assurance regions in DEA. Operations Research 2008;56(1):69–78.
[14] Doyle J, Green R. Data envelopment analysis and multiple criteria decision making. Omega 1993;21(6):713–715.
[15] Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA. Pitfalls and protocols in DEA. European Journal of Operational Research 2001;132 (Issue 2):245–259.
[16] Golany B, Roll Y. An application procedure for DEA. OMEGA 1989;17 (3):237–250.
[17] Liu WB, Meng W, Li XX, Zhang DQ. DEA models with underirable inputs and outputs. Annals of Operations Research 2010;173:177–194.
[18] Liu JS, Lu LYY, Lu W-M, Lin BJY. Data envelopment analysis 1978-2010: a citation-based literature survey. OMEGA 2013;41(Issue 1):3–15.
[19] Paradi JC, Zhu H. A survey on bank branch efficiency and performance research with data envelopment analysis. OMEGA 2013;41:61–79.
[20] Seiford LM, Zhu J. On piecewise loglinear and log efficiency measures. Computers and Operations Research 1998;25(5):389–395.
[21] Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. European Journal of Operational Research 2002;142(1):16–20.
[22] Sherman HD, Joe Zhu. Analyzing performance in service organizations. Sloan Management Review 2013;54(4):37–42 (Summer).
[23] Shwartz M, Burgess Jr. JF, Berlowitz, D. D. Benefit-of-the-doubt approaches for calculating a composite measure of quality. Health Services and Outcomes Research Methodology 2009;9:234–251.
[24] Stewart TJ. Relationships between data envelopment analysis and multicriteria decision analysis. The Journal of the Operational Research Society 1996;47(5):654–665.
[25] Thompson RG, Langemeir LN, Lee C, Lee E, Thrall RM. The role of multiplier bounds in efficiency analysis with application to Kanssa farming. Journal of Econometrics 1990;1(2):93–108.
[26] Tone K. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 2001;130(Issue 3):498–509.
[27] Triantis KP. Engineering applications of data envelopment analysis. In: Cooper WW, Seiford LM, Zhu J, editors. Handbook on data envelopment analysis. New York: Springer; 2011. p. 2011.