Non-metric Data - A Note on a Neglected Problem in DEA Studies

Norbert Hirschauer and Oliver Musshoff (preprint version: November 2013)

The final publication is available at http://link.springer.com (2014, European Journal of Law and Economics 37(3): 489-494; DOI 10.1007/s10657-013-9429-5).

Data Envelopment Analysis (DEA) is widely used to compare the empirical performance of public institutions such as law enforcement agencies, judicial authorities or national health care systems. Many DEA analysts, however, ignore the fact that DEA efficiency values are non-metric. They consequently do not hesitate to compute (arithmetic) means. They do not hesitate either to treat DEA values as metric data in econometric analyses. Instead of providing useful insights into the performance of public bodies, the confusion of non-metric data with metric data constitutes a lack of internal validity that may cause serious fallacies. Against this background, we believe that a clear warning against an uncritical processing and interpretation of DEA values is pertinent and should be routinely considered by efficiency analysts as well as referees of efficiency papers.

Not everything that can be counted [or computed] counts and not everything that counts can be counted (attributed to Albert Einstein).

Keywords Data envelopment analysis; Scales of measurement; Non-metric character of technical efficiency values

JEL Classification C02, C61, H00, K00

Background of efficiency analysis

A major goal of law and economics scholars is to improve the management and the performance of public institutions such as law enforcement agencies, judicial authorities or national health care systems. This applies to input decisions and process management as well as to the adaptation of services towards outcomes that are socially desired. The attempt to contribute to an improved performance of public bodies has prompted researchers to search for "learning opportunities" by means of comparative analysis and benchmarking.

While private sector firms usually evaluate their performance on the basis of profitability indicators, economic criteria reach their limits when outputs and/or inputs cannot be mapped into monetary units. This is frequently the case in the public sector. Accounting for the monetization problem, technical efficiency analyses such as DEA have been widely used to extract information from data sets that provide quantitative but non-monetary information on the resources that are used and the services that are provided by public bodies (cf., e.g., Charnes et al. 1978; Charnes et al. 1981; Charnes et al. 1985; Schneider 2005; Kocher et al. 2006; for an overview of DEA studies see Emrouznejad et al. 2008).

Besides its general meaning in economic theory, the term “efficiency analysis” has become known as a label attached to a specific class of mathematical procedures that are employed for the comparative analysis of the productivity of decision-making units in a sample. Efficiency analysis procedures identify benchmarks units which feature minimum input-output ratios that are understood as forming an empirical production function (frontier) relative to which the position of each decision-making unit is determined. Whilst the quality of the allocation decision cannot be assessed as long as the values (“prices”) of outputs and inputs are not known, it is possible to identify – on this “lowest level of decision-making” (Ray 2004: 14) – at least those units that are inefficient in that they use too much input for a given output (avoidable input waste) or produce too little output for a given input (unrealized output potential).

In his textbook on DEA, Ray (2004: 5) issues the following warning: “At present, an overwhelming majority of practitioners remain content with merely feeding the data into the specialized DEA packages without much thought about whether the LP model solved is really appropriate for the problem under investigation.” Ray’s warning refers to external validity. He emphasizes the need to check models regarding their capacity to answer the research question under consideration. An exemplary challenge would be the question whether using the number of completed court cases as an output measure contributes to our understanding of the relative performance of courts in producing justice; it does so only if we can plausibly assume that there is little variability both in the complexity of cases and the quality of court decisions.

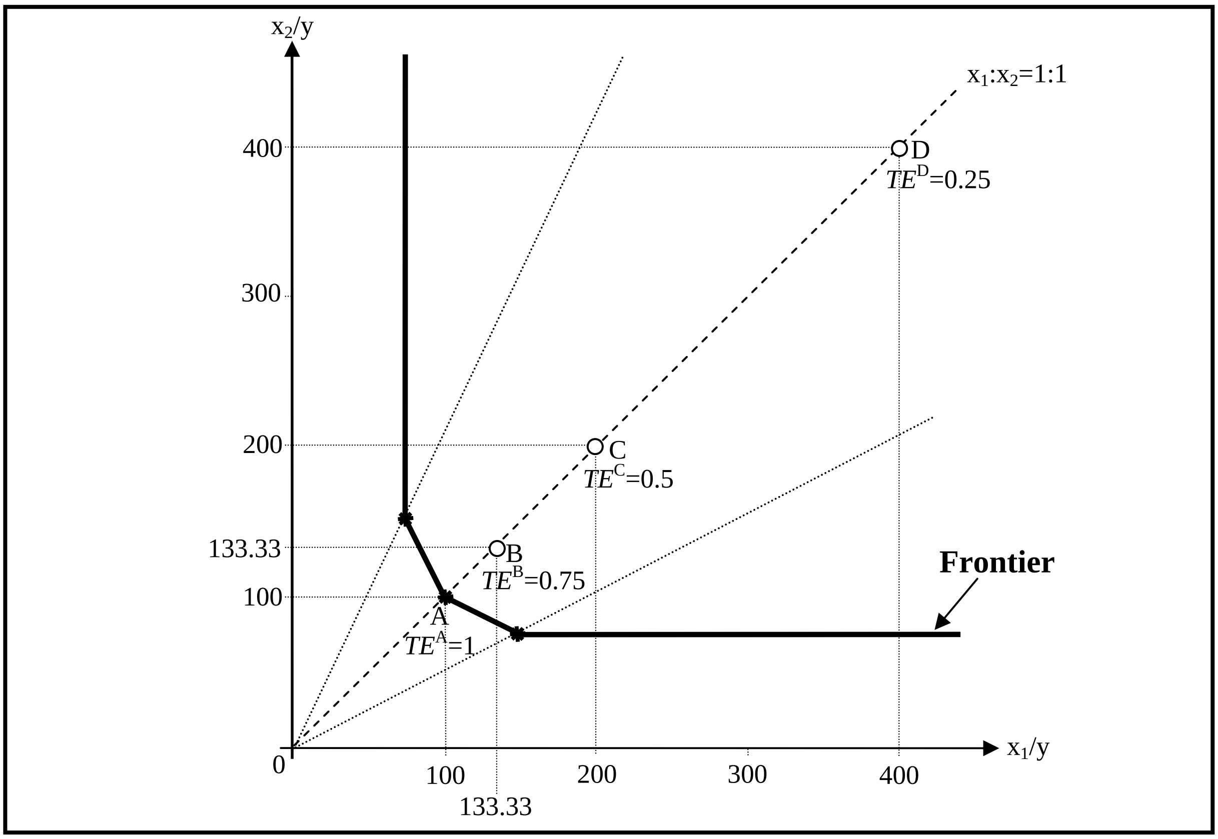
While representing a crucial issue, external validity is not our concern in this note. Instead, we look at a more fundamental problem which arises if the fact is ignored that efficiency values derived from DEA do *not* provide metric information with regard to “substantive” inefficiency, i.e., the avoidable input waste or the unrealized output potential. Confusing non-metric data with metric data jeopardizes internal validity and may reduce the value of a study to less than nothing from the very start of the analysis. It is nonetheless a widespread flaw. We thence believe that an exposition of the level of measurement and a clear warning against an uncritical processing and interpretation of DEA efficiency values is pertinent. In the interest of providing easy intuition how and why DEA values are non-metric, the exposition focuses on a stylized input-oriented DEA model. The non-metric measurement level of DEA values, however, is a feature caused by its specific geometric design. It is thus shared by all its variants including output-oriented DEA, overall (cost or revenue) efficiency, Malmquist index, super efficiency, and stochastic DEA estimators.

DEA revisited

DEA estimates a non-parametric frontier as an "envelope" of the input-output combinations that are observed in the examined sample of decision-making units. At the same time, efficiency measures are calculated based on the geometric distance between the input-output combination of the respective unit and the frontier (cf. FARRELL 1957). DEA transforms partial productivities into a *single* radial efficiency measure (ROSENBERG 1991). For this purpose, a linear program is solved for each of the examined units (cf. COELLI et al. 1998: 133ff.; RAY 2004: 28ff.).

Figure 1 refers to an input-oriented DEA model for two inputs x_1 and x_2 and a single output y. The frontier is formed by linear segments produced by piecewise linear combinations of the empirically observed minimum input-output ratios (here: the three units represented by bold points). Inefficient units above the frontier are represented by simple points.

Figure 1: Technical efficiency values in the 2-input/single-output case



Units on identical rays share an equi-proportionate mix of inputs. Figure 1 focuses on units B, C, and D which all have a ratio of 1:1 between input x_1 and x_2. Unit A, which has the same input ratio, exhibits the minimum input-output ratio. It is thence the benchmark for the increasingly less efficient units B, C, and D. Let us look at unit B to understand the DEA calculus. The so-called "technical efficiency" of B is a radial measure which corresponds graphically to the ratio of the distances $\overline{0A}$ and $\overline{0B}$ (cf. COELLI et al. 1998: 52). The technical efficiency values of units A, C, and D are calculated correspondingly (cf. Table 1).

Table 1: The non-metric character of "technical efficiency" values

Unit	**A**	**B**	**C**	**D**
TE	$\overline{0A}/\overline{0A} = 1.00$	$\overline{0A}/\overline{0B} = 0.75$	$\overline{0A}/\overline{0C} = 0.50$	$\overline{0A}/\overline{0D} = 0.25$
Input 1 Waste 1	100 -	133.33 33.33	200 100	400 300
Input 2 Waste 2	100 -	133.33 33.33	200 100	400 300
Output	1	1	1	1

What DEA users should consider

Table 1 illustrates the efficiency assessment provided by DEA. To those interested in the "true" efficiency of decision-making units within a sample, the question of interest is how much avoidable input waste these units have. According to the DEA calculus, the decrease in efficiency is a constant 0.25 from unit to unit. These equal differences, however, correspond to input wastes that increase from 33.33 in the case of B, to 100 in the case of C, and to 300 in the case of D. This is due to the fact that the DEA calculus changes its point of reference (denominator) from unit to unit. Consequently, DEA values are non-metric data and differences between them are not meaningful with regard to substantive efficiency. That is, an equal distance between DEA values must not be interpreted as an equal distance with regard to the inefficient resource use (i.e., the avoidable input waste).

There is an easy way out: using the reciprocals of the non-metric DEA values (e.g., 2 instead of 0.5, and 4 instead of 0.25 etc.) provides a metric measurement in the form of an index that reflects the inputs used by the unit under consideration compared with the inputs used by its benchmark unit (base value). An index of 2 (4), for instance, informs us that the unit uses twice (four times) the amount of inputs compared to its benchmark. Deducting 1 from the index, we could alternatively formulate that the unit is so inefficient as to have 100% (300%) input waste compared to the input use of its benchmark.

The input wastes of unit B (*TE* = 0.75), unit C (*TE* = 0.5), and unit D (*TE* = 0.25) amount to 33.33%, 100%, and 300%, respectively. The total input waste of these three inefficient units amounts to 433.33% of the inputs used by their benchmark A. Their average input waste is 144.44% (= 433.33% / 3). While this average is meaningful, the 0.5 average of the three corresponding DEA values (0.75, 0.5, 0.25) is not! A DEA value of 0.5 equals an input waste of 100% and conveys the information “being half as efficient as the benchmark”. The three units B, C and D have neither an average input waste of 100% nor are they, on average, half as efficient as their benchmark A. However, impairing internal validity, exactly that “information” is conveyed if one computes the arithmetic mean of DEA values.

In line with the “Theory of Scales of Measurement” (cf. Stevens 1946), which is concerned with the question which mathematical operations are permissible for which measurement level, DEA users should consider the limited information content of DEA values:

- Arithmetic operations (addition/subtraction) and the calculation of arithmetic means require variables to exhibit a metric scale level. These operations are meaningless for non-metric variables such as DEA values.
- Being non-metric implies that the formation of ratios in the sense of statements such as “business X is twice as efficient as business Y” is not permissible either for DEA values.
- While they could be transformed into their metric reciprocals, DEA values by themselves represent at best an ordinal-scaled variable for the units of a given sample. Hence, only the operations “more efficient” (>), “less efficient” (<) or “equally efficient” (=) are permissible within the respective sample.

Unfortunately, DEA values are not commonly transformed into their more easily manageable metric reciprocals even though this is the prerequisite for widespread operations, such as averaging and the treatment as metric data in econometric analyses.

Besides the level of measurement problem, attention should be paid to another critical issue: taking the very best units out of the sample will increase DEA values. In other words: DEA values will be very high even if all units in the sample are substantially inefficient but homogeneously so. It is thence impossible to compare DEA values between samples.

We will finally conclude by hinting at an imminent semantic problem. DEA is, first of all, a mathematical procedure which prescribes a certain way of how to process physical input and output data. Attaching the label "*efficiency*", which is used in economic theory, to the values derived from this procedure entails the risk of misinterpretations. The results of a measurement procedure must never automatically be equated with the actual content of interest – even if identical terms are attached to both. While a critical appraisal in this regard is necessary for all measurements, it is especially relevant in the case of DEA because, due to its non-metric measurement level, the so-called "technical efficiency" conveys very little information regarding the actual content of interest, i.e., the level of the inefficient resource use.

Literature

Charnes, A., Clark, C.T., Cooper, W.W. and B. Golany. 1985. A Developmental Study of Data Envelopment Analysis in Measuring the Efficiency of Maintenance Units in the US Air Forces. *Annals of Operations Research* 2 (1), 95-112.

Charnes, A., Cooper, W. and E. Rhodes. 1978. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research* 2 (6), 429-444.

Charnes, A., Cooper, W. and E. Rhodes. 1981. Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through. *Management Science* 27 (6), 668-697.

Coelli, T., Prasada Rao, D.S. and G.E. Battese. 1998. An Introduction to the Efficiency and Productivity Analysis. Kluwer Academic Publishers, Norwell.

Emrouznejad, A., Parker B.R. and G. Tavares 2008. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences* 42, 151–157.

Farrell, M.J. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A (General)* 120 (3), 253-290.

Kocher, M.G., Luptacik, M. and M. Sutter. 2006. Measuring Productivity of Research in Economics: A Cross-country Study using DEA. *Socio-Economic Planning Sciences* 40 (4), 314-332.

Ray, S.C. 2004. Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research. Cambridge University Press, New York.

Rosenberg, D. 1991. DEA or Dubious Efficiency Assessment? *Journal of Health and Human Resources Administration* 14 (1), 65-76.

Schneider, M. 2005. Judicial Career Incentives and Court Performance: An Empirical Study of the German Labour Courts of Appeal. European Journal of Law and Economics 20, 127-144.

Stevens, S.S. 1946. On the theory of scales of measurement. *Science* 103 (2684), 677-680.