Socio-Economic Planning Sciences 52 (2015) 41–54

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

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

Adjusting for the environment in DEA: A comparison of alternative models based on empirical data

Jean-Marc Huguenin*

Swiss Graduate School of Public Administration, University of Lausanne, Switzerland

A R T I C L E   I N F O

Article history:
Available online 24 October 2015

Keywords:
Data envelopment analysis
Environmental variables
Comparison
Empirical data

A B S T R A C T

Due to the existence of free software and pedagogical guides, the use of Data Envelopment Analysis (DEA) has been further democratized in recent years. Nowadays, it is quite usual for practitioners and decision makers with no or little knowledge in operational research to run their own efficiency analysis. Within DEA, several alternative models allow for an environmental adjustment. Four alternative models, each user-friendly and easily accessible to practitioners and decision makers, are performed using empirical data of 90 primary schools in the State of Geneva, Switzerland. Results show that the majority of alternative models deliver divergent results. From a political and a managerial standpoint, these diverging results could lead to potentially ineffective decisions. As no consensus emerges on the best model to use, practitioners and decision makers may be tempted to select the model that is right for them, in other words, the model that best reflects their own preferences. Further studies should investigate how an appropriate multi-criteria decision analysis method could help decision makers to select the right model.

1. Introduction and objectives

The use of Data Envelopment Analysis (DEA) is experiencing rapid and continuous growth. In 2002, Tavares [67] identified 3203 publications. In 2008, Emrouznejad et al. [27] inventoried more than 7000 publications. This growth reflects the need for user-friendly performance measurement methods. In recent years, the use of DEA has been further democratized due to *(1)* the existence of free software, *(2)* the publication of pedagogical guides [19,43] and *(3)* the teaching of DEA in under- and postgraduate programs.[1] Nowadays, it is quite usual for practitioners and decision makers with little or no background in operational research and economics to run their own efficiency analysis. For instance, a web-based platform integrating DEA has been developed in Portugal for secondary schools' headteachers [57].

The external environment could influence the ability of management to convert inputs into outputs and, as a result, impact entities' technical efficiency. Following Coelli et al. [20]; (p. 190), an environmental variable is defined as a factor that could influence the efficiency of an entity, where such a factor is not a traditional input and is assumed to be outside of the manager's control.

Within DEA, several models allow for an environmental adjustment. There are few published studies which compare these models with one another. The aim of this study is to test how several alternative DEA models can possibly deliver diverging results. Unlike studies using simulated data, this study intentionally uses empirical data. As a result, the comparison is made between the estimates of the alternative models, without knowing whether these estimates approximate the 'true' efficiency measure.[2] By

* University of Lausanne, Swiss Graduate School of Public Administration, Bâtiment IDHEAP, 1015 Lausanne, Switzerland. Tel.: +41 (0)32 941 64 18.
E-mail address: jean-marc.huguenin@unil.ch.

[1] For instance, DEA is taught at the University of Lausanne, Switzerland, in three different courses: *(1)* Public Sector Performance Measurement (Master of Science in Public Policy and Management), *(2)* Public Sector Financial Management (Master of Advanced Studies in Public Administration) and *(3)* Benchmarking (Certificate of Advanced Studies in Administration and Management of Educational Establishments). About 90 decision makers in the public sector are trained annually in the use of DEA.

[2] As a reviewer of this paper points out, another research question would be to determine whether the estimates of alternative models converge or diverge with the 'true' efficiency. This question cannot be answered by using empirical data, as the 'true' efficiency is unknown. To answer this question, the 'true' efficiency would have to be estimated by (1) defining a production function, (2) generating inputs from a random distribution and (3) deriving the output. This would be technically feasible. However, this study positions itself from the standpoint of practitioners and decision makers with no or little knowledge in operational research or economics. The additional step of considering the 'true' efficiency has not been retained in this study because it introduces a supplementary difficulty and sophistication for practitioners. The difficulty of generating a data set with the same patterns of the empirical data considered in this study, in addition to the fact that multiple outputs are considered, appears probably too demanding for practitioners. Note also that existing studies using simulated data provide mixed results about the convergence of alternative models with the 'true' efficiency (see Section 2.2.1 about it).

http://dx.doi.org/10.1016/j.seps.2015.10.004
0038-0121/

using empirical data, this study addresses the issue faced by practitioners and decision makers who perform their own efficiency analysis. It seeks to determine whether the alternative models produce convergent results (i.e. consistent efficiency scores and rankings of entities). If the alternative models do produce consistent results, practitioners and decision makers may confidently select any model. If they produce divergent results, the choice of model becomes a strategic issue.

The alternative models tested in this study are all user-friendly and easily accessible to practitioners and decision makers. The empirical case is the 90 primary schools of the State of Geneva, Switzerland. It is particularly well suited to test several alternative models, as *(1)* the State of Geneva practices positive discrimination towards disadvantaged schools and *(2)* schools are grouped in five categories defined by one continuous variable reflecting the economic status of pupils. According to their respective category, schools receive additional teaching staff.

2. Literature review

2.1. Adjusting for the environment in DEA

DEA is an efficiency measurement technique. It was originally developed by Ref. [74]. A comprehensive treatment of the methodology is proposed by Cooper et al. [21].

Within DEA, several models allow for an environmental adjustment. Following Muñiz [54]; they can be grouped into three categories: *(1)* one-stage models [7,8,59]; Yang and Paradi model in Ref. [55], *(2)* multi-stage models including two-stage [58], three-stage [33,54,60] and four-stage models [34] and *(3)* program analysis models [17]. This taxonomy can be completed by adding the conditional nonparametric models (see Ref. [6]; for a review of conditional nonparametric models). These models can be applied to full or partial order frontiers (such as order-*m* frontiers or order-α quantile type frontiers). The basic model of this category has been developed by Ref. [24] based on the preliminary work of [18]. In Ref. [24]; the conditional efficiency measures are defined and estimated nonparametrically. The most recent model of this category is provided by Ref. [6]. Finally and based on the concept of metafrontier function introduced by Ref. [40]; metafrontiers models using the DEA technique have been developed [56]. These models are appropriate when dealing with decision-making units from different industries, regions or countries. The concept of Malmquist productivity index is an extension of metafrontier models to panel data [31].

Among the models which allow for an environmental adjustment, four of them are retained in this study because they are all, to some extent, user-friendly and easily accessible to practitioners and decision makers.

The Banker and Morey [7] model (BM1986a) and the Banker and Morey [8] model (BM1986b) are the first two models retained. In BM1986a, the decision making units (DMUs) are grouped into homogenous categories. These categories are defined by the level of the environmental variables. In order to calculate efficiency, DMUs are compared only with other DMUs with similar or worse environmental variables. In BM1986b, the environmental variables are included directly into the model as non-discretionary variables. This model takes into account the fact that environmental variables are not under the control of management and cannot be treated as discretionary factors. As a result, the constraints on the environmental variables are modified. Interested readers will find the specification of these models in Refs. [7,8].

Although they have been criticized, Harrison, Rouse and Armstrong [39] note that these models are widely used by researchers. They have generated at least 239 different publications [50]. Harrison et al. [39] mention that it suggests that many researchers have found these models appropriate for their particular context. They also mention that "given there is no DEA model that is clearly superior in controlling for non-discretionary inputs, researchers continue to refer to the work of [7,8] (p. 263)". See for instance [35] for an application of BM1986a and [54] for an application of BM1986b.

The third model retained is the Ray [58] model (R1991). This model contains two stages. In the first stage, a basic DEA model is performed using only discretionary variables. After obtaining the technical efficiency scores (TE) from the first stage, R1991 uses an ordinary least squares model to regress these scores upon environmental variables in the second stage. Since Ray [58]; other types of regression have been used in the second stage. For instance, McCarty and Yaisawarng [51] are the first to use a Tobit regression. R1991 is recommended by Ref. [20] in most cases. It has demonstrated its superiority to other models which allow for an environmental adjustment [60,62]. See for instance Burney et al. [14] for an application of R1991. Interested readers will find the specification of this model in Ref. [58]. Note that Simar and Wilson [65] have showed that the results of the second-stage regression could be biased and invalid unless some restrictive conditions are fulfilled. These authors proposed a complex seven-step estimation procedure which includes a double bootstrapping to overcome this bias. However, more recent studies have showed that OLS may actually provide consistent estimates in the second-stage regression [9,52]. McDonald [52] also proves that, if the White correction is applied to the model to correct for heteroskedasticity [69], sample tests can be performed which are robust to the distribution of the disturbances.

Finally, the fourth model retained is the Yang and Paradi model in Ref. [55]; (p. 1176) (YP2006). This model applies a handicapping measure based on the levels of the non-discretionary variables. Entities with a favourable environment are penalized by the handicapping measure. Non-discretionary inputs are adjusted with a higher handicap and non-discretionary outputs are adjusted with a lower handicap. As a result, adjusted inputs have a higher value than original inputs and adjusted outputs have a lower value than original outputs. Interested readers will find the specification of YP2006 in Ref. [55].

YP2006 is relatively little known and used. Compared to BM1986a, it does not lessen the discriminating power of DEA, as it does not categorize the entities. YP2006 is particularly suited when discretionary inputs and/or outputs are augmented or diminished according to the condition of the environment. See for instance Ref. [71] for an application.

2.2. Comparing the models

Various studies have conducted benchmark analysis of alternative methods to measure efficiency (such as corrected ordinary least squares, stochastic frontier analysis, DEA or Free Disposal Hull). Evidence suggests that the choice of technique affects efficiency scores and rankings of entities. See Ref. [47]; (pp. 661–662) for a short review. For instance, Farsi and Filippini [32] assess the electricity distribution utilities in Switzerland. They study the sensitivity of three benchmarking methods, one being nonparametric (DEA) and two being parametric. Their results indicate that both efficiency scores and rankings of entities are significantly different across methods.

Alternative models to measure efficiency, within DEA, can also lead to diverging results but this has been far less investigated. Whilst few studies address this issue, interest seems to have been growing in recent years.

Some studies [23,30,39,55,59,60,62] use simulated data. The objective of these studies is to allow for comparisons between