Applied Soft Computing 46 (2016) 204–219

Contents lists available at ScienceDirect

Applied Soft Computing

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

An integrated data envelopment analysis and free disposal hull framework for cost-efficiency measurement using rough sets

Rashed Khanjani Shiraz[a], Hirofumi Fukuyama[b], Madjid Tavana[c,d,*], Debora Di Caprio[e,f]

[a] School of Mathematics Science, University of Tabriz, Tabriz, Iran
[b] Faculty of Commerce, Fukuoka University, Fukuoka, Japan
[c] Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, PA 19141, USA
[d] Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, D-33098 Paderborn, Germany
[e] Department of Mathematics and Statistics, York University, Toronto M3J 1P3, Canada
[f] Polo Tecnologico IISS G. Galilei, Via Cadorna 14, 39100 Bolzano, Italy

ARTICLE INFO

Article history:
Received 9 November 2014
Received in revised form 8 April 2016
Accepted 30 April 2016
Available online 6 May 2016

Keywords:
Data envelopment analysis (DEA)
Free disposal hull (FDH)
Rough sets
Cost-efficiency

ABSTRACT

Traditional cost-efficiency analysis methods require exact and precise values for inputs, outputs and input prices. However, this is not the case in many real-life applications. This study proposes a rough cost-efficiency approach to the problem of ranking efficient decision making units (DMUs). Based on rough set theory, a nonparametric methodology for cost-efficiency analysis is developed. The merits of this methodology include computational ease and the capacity to incorporate data uncertainty. Furthermore, it applies to both convex data envelopment analysis (DEA) and non-convex free disposal hull (FDH) technologies under different returns-to-scale assumptions. A numerical example and a real-life case study in the Japanese banking industry demonstrate the applicability of the proposed framework. In particular, the rankings of the DMUs resulting from the proposed models are compared with those obtained using the maximum technical efficiency loss index.

1. Introduction

In this study, we develop two cost-efficiency models based on data envelopment analysis (DEA) and free disposal hull (FDH), respectively. The purpose of the proposed models is to provide a method to evaluate the cost-efficiency of decision-making units (DMUs) when there exists a significant level of impreciseness in the data while allowing for different returns-to-scale assumptions.

1.1. Traditional DEA and FDH models versus imprecise data

The traditional DEA model, initially proposed by Charnes et al. [7] and extended by Banker et al. [3], is a nonparametric mathematical programming method for evaluating the relative efficiency of decision-making units characterized by crisp multiple inputs and outputs. It consists of solving a fractional linear programming problem through an equivalent linear programming formulation assuming convexity and constant returns-to-scale (CRS).

The traditional DEA model has received considerable attention in both theory and applications since the very beginning (see Ref. [18] or Ref. [12] for a comprehensive bibliography) quickly becoming an important research tool in management science, operations research, and decision theory. Regarding, in particular, the study of cost-efficiency and DMUs' performances, Färe et al. [15] operationalized Farrell's [16] cost-efficiency notion. This cost-efficiency measure requires operating with crisp input and output data.

---

[*] Corresponding author at: Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, PA 19141, USA. Fax: +1 2672952854.

E-mail addresses: Rashed.shiraz@gmail.com (R. Khanjani Shiraz), fukuyama@fukuoka-u.ac.jp (H. Fukuyama), tavana@lasalle.edu (M. Tavana), dicaper@mathstat.yorku.ca, debora.dicaprio@istruzione.it (D. Di Caprio).

http://dx.doi.org/10.1016/j.asoc.2016.04.043

The traditional FDH model, proposed by Deprins et al. [8], is an other nonparametric deterministic method useful to evaluate the technical efficiency of DMUs. It exploits the input–output disposability without imposing the convexity assumption. Among others, Tulkens [52] presented a FDH-based mixed integer linear programming while Kerstens and Vanden Eeckaut [27] introduced various returns-to-scale specifications, namely, non-increasing, non-decreasing, and constant returns-to-scale. Agrell and Tind [1] suggested a linear programming FDH model that was later extended by Leleu [30] to various returns-to-scale technologies using cost functions.

Both the traditional DEA and FDH models require the values of inputs and outputs to be known precisely. However, quite often, the observed data are imprecise, vague or taken at one point in time, which may not reflect the overall distribution of the data (such as, for example, the total working hours of employees). In this regard, Hougaard [21] stated three reasons due to which the traditional DEA model is inadequate to model some real life situations and evaluate the relative efficiency scores. First, the efficiency scores are very sensitive to changes in the data and to errors in the estimation of the efficient frontier of the technology. Second, the relative quality differences among decision-making units (DMUs) in terms of inputs and outputs, may distort the true efficiency performance landscape. Third, efficiency scores are simply approximations of the DMU's unknown preferences. The traditional FDH model shows the same limitations.

1.2. Main shortcomings of interval models with imprecise data

Many performance measurement techniques based on interval values have been proposed to account for the impreciseness and vagueness of data relative to production technologies.

Entani et al. [14] suggested evaluating the efficiency of a DMU by an interval delimited by a pessimistic and an optimistic measure. Despotis and Smirlis [9] introduced an interval DEA method to deal with imprecise data. Extending Despotis and Smirlis' [9] method, Kao [26] constructed a two-level mathematical programming model to facilitate the calculation of efficiency intervals for ordinal data. Kao [26] also supported the interval approach from a psychological viewpoint arguing that DMUs are willing to accept interval measures better than crisp ones since the former ones do not directly imply that the performance of a DMU is worse than that of others. Inuiguchi and Mizoshita [22] examined some DEA models with interval input-output data discussing how to obtain lower and upper bounds for the efficiency scores.

Although frequently used to model real-life problems, the interval models and the corresponding solution methods may not be versatile enough to provide a satisfactory evaluation of DMUs such as banks or manufacturing firms.

Three different approaches have been suggested through the years for handling vagueness and impreciseness in DEA models: (1) stochastic, (2) fuzzy, and (3) rough.

The stochastic DEA approach makes it possible to replace crisp data with statistical or probabilistic values. In this approach, chance-constrained formulations are usually introduced and the uncertainty relative to the available data incorporated by interpreting inputs and outputs as random variables whose cumulative distribution functions and probability density functions are known. However, solving stochastic DEA with various cumulative and probability distributions is often very complex from the computational and, hence, practical viewpoint.

The fuzzy DEA approach uses membership functions to model uncertainty and imprecise data sets while the rough DEA approach centers on lower and upper approximations of crisp sets.

All the DEA approaches described above focus on constructing interval values. Note though that, properly speaking, only stochastic and fuzzy DEA are interval methods since they require the existence of lower and upper bounds of inputs and outputs. On the other hand, the rough DEA approach allows for a more general interpretation of data uncertainty since rough variables are used to define lower and the upper approximations of the interval values corresponding to the available inputs and outputs.

Hence, the main shortcoming of interval models, that is, the fact that they can be applied only to situations where inputs and outputs are already endowed with both a lower and an upper bound, is directly inherited from stochastic and fuzzy DEA while it remains a tangential issue in rough DEA.

Besides the aforementioned shortcomings, rough DEA also has the same limitations of a classical DEA model, that is, it applies only to convex technologies and usually assumes constant returns-to-scale. Thus, in order to complement our analysis and extend the efficiency evaluation problem to the non-convex case, a rough FDH approach is necessary to integrate the DEA framework. Moreover, a rough FDH model also allows for considering different returns-to-scale assumptions.

1.3. Contribution

In summary, there already exists a very ample literature on DEA and FDH, most of which accounting for the uncertainty characterizing the inputs and outputs of many real-life situations where a measure of the relative efficiencies of a set of DMUs is required. In particular, the flexibility of the DEA approach has been exploited to solve a myriad of input–output efficiency evaluation models, including chance-constrained ones, whose inputs and outputs are fuzzy, random, rough, random-rough or fuzzy-rough variables.

All the models proposed in the previous studies are at some extend based on the interval approach and each of them presents difficulties and advantages when dealing with specific situations.

The current study focuses on:

- Developing a cost-efficiency model for convex technologies able to incorporate the impreciseness of inputs, outputs and input prices within the DEA framework while reducing the computational complexity of the traditional approach;
- Developing a rough FDH approach to the cost-efficiency evaluation problem for non-convex technologies so as to complement rough DEA cost-efficiency measures when non-constant returns-to-scale are assumed.

An attempt to propose rough DEA as the most appropriate framework to deal with efficiency evaluation problems was made by Xu et al. [54] who developed a DEA model with rough parameters to evaluate the performance of real supply chains. However, to the best of