European Journal of Operational Research 254 (2016) 946–956

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

DEA models for non-homogeneous DMUs with different input configurations

WangHong Li[a], Liang Liang[a], Wade D. Cook[b,*], Joe Zhu[c,d]

[a] School of Management, University of Science and Technology of China, 230026 He Fei, An Hui Province, PR China
[b] Schulich School of Business, York University, 4700 Keele Street, M3J 1P3 Toronto, Ontario, Canada
[c] International Center for Auditing and Evaluation, Nanjing Audit University, Nanjing, PR China
[d] School of Business, Worcester Polytechnic Institute, Worcester, USA

ARTICLE INFO

Article history:
Received 1 December 2015
Accepted 28 April 2016
Available online 6 May 2016

Keywords:
Data envelopment analysis
Non-homogeneous DMUs
Missing inputs
Multiple processes
Input configurations

ABSTRACT

The data envelopment analysis (DEA) methodology is a benchmarking tool where it is generally assumed that decision making units (DMUs) constitute a homogeneous set; specifically, it is assumed that all DMUs have a common (input, output) bundle. In earlier work by the authors the issue of non-homogeneity on the output side was investigated. There we examined a set of steel fabrication plants where not all plants produced the same set of products/outputs. In the current research we investigate non-homogeneity on the input side. Such can occur in manufacturing plants, for example, when the output bundle can be produced using different mixes of machines, robots and laborers. Thus, we can have an input *configuration* existing in a DMU that is different from the configuration in another DMU. As a practical application of this phenomenon, we examine the measurement of efficiencies of a set of provinces in China. There, all provinces have the same common set of outputs in the form of GDP, supported population, and an undesirable output, nitrogen dioxide. On the input side, however, this commonality is missing. While all provinces have water, capital investment and natural resources, the latter of these (natural resources) takes several different forms, namely coal, natural gas and petroleum. However, not all provinces have the same mix of these resources, nor are there clear exchange rates among these very different, albeit substitutable inputs. This means that that one cannot directly apply the conventional DEA methodology. This then raises the question as to how to fairly evaluate efficiency when the configuration or mix of inputs can differ from one DMU to another. To address this, we view the generation of outputs for a province as a set of processes created by the different configurations of natural resources available. We develop a DEA type of methodology to evaluate these processes. This evaluation provides important insights into not only the overall performance of each province, but as well provides measures of the efficiency of the various configurations of the three natural resources.

1. Introduction

Data Envelopment Analysis (DEA), first introduced by Charnes, Cooper, and Rhodes (1978), is a methodology for evaluating the relative efficiencies of a set of decision making units (DMUs). In the nearly four decades since this seminal work, literally thousands of articles and books on DEA have appeared. Useful surveys include Cook and Seiford (2009) and Paradi and Zhu (2013). The conventional DEA model is based on the assumption that in a multiple-input multiple-output setting, all inputs impact all outputs, or that detailed input to output relations are not closely examined. Furthermore, it is assumed in the conventional DEA model that the set of DMUs under investigation constitute a homogeneous set. This means that all DMUs have the same inputs and produce the same outputs.

There are many situations in which the above assumptions are violated. Regarding the assumption that all inputs affect all outputs, consider the situation in a manufacturing setting where one of the inputs is packaging resources. Clearly, this input only impacts outputs that require packaging. This gives rise to what we refer to as partial input to output impacts, investigated in Cook, Imanirad, and Zhu (2013a). On the matter of the DMUs constituting a homogeneous set, consider the case in manufacturing where some DMUs may produce a different output mix than is true of other DMUs. See Cook, Harrison, Rouse, and Zhu (2012), Cook, Harrison, Imanirad, Rouse, and Zhu (2013b), where the efficiency of a

* Corresponding author. Tel.: +1 416 736 2100/33573.
E-mail address: wcook@schulich.yorku.ca (W.D. Cook).

http://dx.doi.org/10.1016/j.ejor.2016.04.063
0377-2217/

set of steel fabrication plants is examined. This latter is a form of non-homogeneity on the *output* side. It was assumed there that the input set was common across all DMUs.

In another setting, where one is comparing a set of universities, for example, it can occur that not all institutions have the same departments, or the same financial resources to attract students. This might be deemed a case where there is lack of homogeneity on the *input* side. In the current paper we extend that earlier research of Cook et al. (2012, 2013b) to encompass the case where the *input mix* can be different for some DMUs as compared to others.

In Section 2 we present a problem setting in which a set of macro-level inputs are used to generate a set of macro-level outputs; one of those inputs is the quantity of natural resources available to the decision making unit. In the case where the DMUs are a set of regions in China, those resources can take different forms in some regions versus other regions. For example, some regions have natural gas while others do not. At the same time, all regions possess natural resources in the form of coal. Hence, the conventional DEA model cannot handle this particular situation where multiple forms of an input occur. In Section 3 we view the multiple forms of an input from the perspective of a set of parallel processes. Section 4 develops a type of DEA-based model to capture this situation, and Section 5 applies this new methodology to the setting discussed in Section 2. Discussion and recommendations appear in Section 6.

2. On regional resources in China

With three decades of rapid economic development in China, people more and more realize that developments in science and technology have promoted productivity and enlarged the Chinese economy. While that is important, we need to realize that ecological preservation is an important core value in Chinese culture. On June 27, 2015, at an important forum, China's Ecological Civilization Guiyang International Forum, the Chinese President set important guidelines regarding the ecology of China. At the same time, the Dean of the Chinese Academy of Sciences, in a speech at the meeting, stressed the importance of scientific research focused on the "optimization of national spatial development patterns, comprehensively promoting resource conservation, the natural ecosystem and environmental protection, and strengthening the construction of the ecological civilization system."

A number of studies have examined various aspects of energy economics in the presence of pollution and environmental concerns. See for example Cooper, Huang, Li, Lelas, and Sullivan (1996), Lee (2005), Soytas and Sari (2009), Zhou, Ang, and Poh (2008), and Zhou and Ang (2008). In the current paper we evaluate the efficiency of a set of regions at a macro-level in terms of the economy and the environment in China; thirty-one regions are examined. Data was collected on five input factors, namely local water resources, capital investment, and natural resources in the form of coal, natural gas and petroleum. On the output side, two important factors to consider for use in evaluating regional economic efficiency are regional GDP and Population served. While one might postulate utilizing these economic factors as two separate outputs in a DEA analysis, a more viable choice is to combine them into a single factor, namely GDP per capita (i.e., GDP/Population). In so doing, proportional increases in this combined factor in inefficient DMUs would represent improvements in the wellbeing of society. If one were, instead, to model the two factors as separate outputs, say in an output oriented setting, an inefficient DMU could become efficient only by proportionally increasing *each* factor so as to project them onto the efficient frontier. There are two problems with this line of reasoning. First, logic would normally dictate that Population be viewed as nondiscretionary, since one cannot simply increase population in a given region. Second, if hypothetically Population were viewed as discretionary, then the same proportional increase in the two variables to get to the frontier would mean that the frontier projection would see GDP per capita at the same level after the projection as before. Such a conclusion would appear to undermine the intended purpose of efficiency analysis.

In developing the regional economy, pollution is a major issue as well; Carbon Dioxide, Nitrogen Dioxide, Sulfur Dioxide, Ammonia, Nitrogen, Smoke and Dust are a natural consequence when a region sets out to improve GDP and the feeding population. Those undesirable outputs pollute the environment; in some areas, the ecological environment can be destroyed. For example, since 2013, "fog" and "haze" have become the annual keywords; Nitrogen Dioxide and Sulfur Dioxide at a level of PM2.5 were major contributors to the hazy atmosphere. In recognition of the importance of pollution and its impacts, Nitrogen Dioxide was selected as one of the most undesirable outputs impacting efficiency in terms of the environment.

Not all regions have the same natural resources. The reality is that some are missing certain forms of this macro input, or more generally, the configuration of inputs differs among the regions. Specifically, the 31 regions fall into 4 groups according to their natural resource "holdings": Regions in the first group have only coal, the second group, coal and petroleum, the third, coal and natural gas, and finally regions in the fourth group have all three forms. All regions hold water resources and investments. This is illustrated by the data in Table 1. This data is published by the China Statistical Yearbook (2013) and China Statistical Yearbook on the Environment (2013).

Utilizing this backdrop, we set out to extend earlier research by Cook et al., (2012, 2013b), by developing a model structure to handle this missing input situation.

3. The problem of multiple configurations of inputs in DEA

The conventional Data Envelopment Analysis model (output oriented CCR) takes the form:

$$\begin{aligned} &\min \frac{\sum_i \nu_i x_{ij_o}}{\sum_r u_r y_{rj_o}} \\ &subject\ to \\ &\quad \frac{\sum_i \nu_i x_{ij}}{\sum_r u_r y_{rj}} \geq 1, \quad j = 1, \ldots n \\ &\quad \nu_i,\ u_r \geq 0 \end{aligned} \tag{3.1}$$

Here, the constraints require that multipliers v_i, u_r be chosen such as to insure that the efficiency ratio of inputs to outputs for each DMU*j* be at or greater than unity. In the conventional DEA model it is assumed that a common set of inputs and outputs characterize all decision-making units (DMUs).

Consider the problem discussed in Section 2 where some DMUs (provinces, cities or regions) have only coal as a natural resource, while others have that resource as well as natural gas, and still others have petroleum as well. Referring to that particular problem setting, let j = 1, 2,..., n denote the DMUs Let x_{1j}, x_{2j} denote the quantities of two inputs, water and capital investment, respectively, available to $DMU_{j.}$ These two inputs are held (in different quantities) by all DMUs. Three additional inputs in the form of natural resources are x_{3j}(coal), x_{4j}(natural gas) and x_{5j}(petroleum), are held in varying amounts by different DMUs. Let y_{1j}, y_{2j} denote outputs GDP per capita, and an undesirable output, $N0_2$. We assume each DMU possesses both of these outputs.

What is different about this problem setting, as indicated above, is that while coal x_{3j} is common to all DMUs, this is not case for x_{4j} and x_{5j}. Fig. 1 shows the DMUs split into 4 groups according to the combination of the three natural resources held by those DMUs.