European Journal of Operational Research 248 (2016) 725–730

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Introducing and modeling inefficiency contributions

Mette Asmild[a,*], Dorte Kronborg[b], Kent Matthews[c]

[a] Department of Food and Resource Economics, University of Copenhagen, Denmark
[b] Center for Statistics, Department of Finance, Copenhagen Business School, Denmark
[c] Cardiff Business School, University of Cardiff, United Kingdom

ARTICLE INFO

Article history:
Received 20 November 2014
Accepted 30 July 2015
Available online 6 August 2015

Keywords:
Data Envelopment Analysis (DEA)
Productivity and competitiveness
Multi-directional Efficiency Analysis (MEA)
Directional data
Chinese banks

ABSTRACT

Whilst Data Envelopment Analysis (DEA) is the most commonly used non-parametric benchmarking approach, the interpretation and application of DEA results can be limited by the fact that radial improvement potentials are identified across variables. In contrast, Multi-directional Efficiency Analysis (MEA) facilitates analysis of the nature and structure of the inefficiencies estimated relative to variable-specific improvement potentials.

This paper introduces a novel method for utilizing the additional information available in MEA. The distinguishing feature of our proposed method is that it enables analysis of differences in inefficiency patterns between subgroups. Identifying differences, in terms of which variables the inefficiency is mainly located on, can provide management or regulators with important insights. The patterns within the inefficiencies are represented by so-called inefficiency contributions, which are defined as the relative contributions from specific variables to the overall levels of inefficiencies. A statistical model for distinguishing the inefficiency contributions between subgroups is proposed and the method is illustrated on a data set on Chinese banks.

1. Introduction

The location of observations within the production possibility set provides potentially valuable information about the underlying structure of the production units. In standard efficiency analysis, like Data Envelopment Analysis (DEA) (see Charnes, Cooper, & Rhodes, 1978), only a small part of this information is utilized, specifically each observation's radial distance to the estimated efficient frontier and the corresponding benchmark (including which observations define the benchmark and the slope of the corresponding facet).

When restricting oneself to considering only an observation's radial projection onto the estimated efficient frontier, most of the available information regarding the location and shape of the efficient frontier is disregarded. Any projection of an observation onto a point on the efficient frontier that dominates the observation in question will result in a Pareto improvement. Therefore one could argue that the location (position) of each observation relative to the whole section of the frontier dominating it is important. For example, if the observations' distances to the frontier in one direction generally are much larger than the distances to the frontier in other dimensions, then this pattern in where the (non-radial) inefficiency is located might provide important insights. Imagine, for example, a situation with two inputs being doctors and nurses working in a hospital. If the inefficiency on nurses is larger than that on doctors, then this could indicate that the nurses have more bargaining power which have resulted in additional nursing staff being allocated. Other reasons for having, or allowing, more inefficiency on some input variables as opposed to others could be i) that inefficiency (slack) on the more flexible resources provides spare capacity that functions as a buffer against uncertain demand, ii) that management focus has been on reducing slack on the more expensive resources or on those where their effort and contributions are easiest to measure, or iii) that investments in future performance may appear as inefficiency on capital in the short run (see Asmild, Bogetoft, & Hougaard, 2013).

The use of Multi-directional Efficiency Analysis (MEA) (see Bogetoft & Hougaard, 1999; Asmild, Hougaard, Kronborg, & Kvist, 2003) enables a consideration of the patterns of the inefficiencies. This is done by first identifying the improvement potential in each dimension separately, which provides an ideal reference point. Next, the overall inefficiency is estimated by projecting the observation onto the frontier in the direction of the ideal reference point, resulting in different relative inefficiencies on the various variables, reflecting whether the observations are located closer to the frontier in some dimensions than in others.

In an illustrative example in the following, we consider a case of Chinese banks. Following the established literature that links

* Corresponding author. Tel.: +45 3533 6886.
E-mail address: meas@ifro.ku.dk (M. Asmild).

http://dx.doi.org/10.1016/j.ejor.2015.07.060
0377-2217/

government ownership of banks and weak economic development and low bank efficiency (Barth, Caprio, & Levine, 2001; La Porta, Lopez-de Silanes, & Shleifer, 2002), one might hypothesize that state-owned banks in China have relatively more inefficiency on non-performing loans than joint-stock banks (Asmild & Matthews, 2012). For further evidence see also Cornett, Guo, Khaksari, and Tehranian (2010). Similarly, the evidence from the public choice literature suggest that state-owned enterprizes, including banks, have been used to finance politically motivated projects as well as over-staffing (Megginson, 2005). When comparing production plans between distinct subgroups it is often not sufficient to simply consider the sizes of the absolute inefficiencies in the specific dimensions for the groups. For the case of Chinese banks, it is well-known that the overall level of inefficiency is generally higher for the state-owned banks than for the joint-stock banks (as is found to be the case in most comparisons of public vs. private organizations). That the state-owned banks are generally located further away from the frontier than the joint-stock banks, implies that the inefficiencies on all variables will generally be larger for the state-owned banks. So in order to specifically investigate whether state-owned banks have more inefficiency on labour relative to the inefficiency on other variables (or, in other words, whether a larger portion of their overall inefficiency comes from labour), than the joint-stock banks have, we need to consider not just the absolute inefficiencies on labour but also its relationship to the overall inefficiency. The ratio of the inefficiency on a specific variable relative to the overall inefficiency is in the following referred to as the inefficiency contribution from the variable in question.

In order to compare the inefficiency contributions between subgroups, all the characteristics determining the distributions of the inefficiency contributions, e.g. both the levels and the variations, might be of interest. In terms of examining differences in variations, one could imagine hypotheses along the lines of whether one subgroup has more variation in the inefficiency contribution from a specific variable than another. Returning again to the example of comparing public and private organizations one might, for example, expect a more formulaic production in public organizations (e.g. requiring 10 nurses per doctor), resulting in similar inefficiency patterns as well, and more variation in the private production. Thus one could formulate hypotheses related to differences in variations between subgroups. Concerning differences in levels, the corresponding hypotheses are straightforward, like the previous example of whether state-owned banks in China have a relatively higher inefficiency contribution from e.g. labour than other bank types have. This is one of the issues that will be investigated in the empirical illustration.

In this paper we propose a model that can be used to analyze the inefficiency contributions. As will be evident in the next section, an inefficiency contribution can be expressed as (the cosine of) an angle, and thus it becomes natural to use models for directional statistics, like the so-called von Mises–Fisher distribution (Mardia, 1975).

The rest of the paper unfolds as follows: In Section 2 we define inefficiency contributions and propose a statistical model for their analysis. Section 3 presents an empirical illustration of the proposed method on an empirical case of Chinese banks and, finally, a discussion of the method is provided in Section 4.

2. Methodology

Consider a set of production plans where m inputs, $\mathbf{x} \in \mathbb{R}^m_+$, are used to produce s outputs, $\mathbf{y} \in \mathbb{R}^s_+$ and denote by $\mathbf{z}$ the vector of throughputs (or netputs)[1], $\mathbf{z} = (-\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_- \times \mathbb{R}^s_+$. Let the production possibility set $\mathcal{P}$ be given by $\mathcal{P} = \{(-\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \text{ can produce } \mathbf{y}\}$. It is assumed that $\mathcal{P}$ satisfies standard assumptions of convexity and free disposability. The technology can, for example, be estimated by the use of DEA as shown in Section 2.2 below.

Multi-directional Efficiency Analysis (MEA) is used as the starting point for defining inefficiency contributions. Generalizing the definitions from Bogetoft and Hougaard (1999) to considering adjustments in both inputs and outputs, let the coordinates of the ideal reference point $\mathbf{z}^{oI}$ for some production plan $\mathbf{z}^o = (-\mathbf{x}^o, \mathbf{y}^o)$ be defined as $z_i^{oI} = \max\{z_i' \in \mathbb{R} \mid (z_i', \mathbf{z}^o_{-i}) \in \mathcal{P}\}$, with $-i$=$\{1, \ldots, i-1, i+1, \ldots, m+s\}$. A benchmark is selected, in the direction of the ideal reference point, by first finding $\beta^* = \max\{\beta \in \mathbb{R}^+ \mid \mathbf{z}^o + \beta(\mathbf{z}^{oI} - \mathbf{z}^o) \in \mathcal{P}\}$. Next, the benchmark is given by $\mathbf{z}^{oB} = \mathbf{z}^o + \beta^*(\mathbf{z}^{oI} - \mathbf{z}^o)$.

2.1. Inefficiency contributions

A measure of overall inefficiency is given by the length of the vector $\mathbf{z}^{oB} - \mathbf{z}^o$. Let the inefficiency contribution from dimension l be given as the dimension specific inefficiency, $z_l^{oB} - z_l^o$ relative to the overall inefficiency. Note that the relationship between the lengths of the vectors $\mathbf{z}_l^{oB} - \mathbf{z}_l^o$ and $\mathbf{z}^{oB} - \mathbf{z}^o$ is identical to the relationship between the lengths of the vectors $\mathbf{z}_l^{oI} - \mathbf{z}_l^o$ and $\mathbf{z}^{oI} - \mathbf{z}^o$. For simplicity the latter is used in what follows, in effect making the identification of the benchmark in MEA superfluous.

Denote by $\mathbf{d}^o = (z^{oI} - \mathbf{z}^o) \in \mathbb{R}_+^{p=m+s}$ the vector between a production plan and its ideal reference point. Thus $\mathbf{d}^o$ is the diagonal of a p-dimensional box (hyperrectangle) with the coordinates of $\mathbf{d}^o$=$(d_1^o, \ldots, d_p^o)$ as side lengths. The coordinates of $\mathbf{d}^o$ in the p-dimensional Euclidean space can, for inefficient units, be converted to p hyperspherical coordinates, consisting of $\|\mathbf{d}^o\| = \sqrt{(d_1^o)^2 + \ldots + (d_p^o)^2}$ (the length of the diagonal in the box) and $p-1$ angular coordinates. Let $\mathbf{d}_i^o$ be the vector $(0, \ldots, d_i^o, \ldots, 0)$ of length p. The relation θ_i^o=$\arccos(\mathbf{d}_i^o \cdot \mathbf{d}^{o^{\mathrm{T}}} / \|\mathbf{d}_i^o\| \|\mathbf{d}^o\|)$= $\arccos(d_i^o / \|\mathbf{d}^o\|)$ ($\theta_i^o \in [0, \frac{\pi}{2}]$) converts the inefficiency contribution for a given dimension to a polar angle (the angle between the vector from $\mathbf{z}^{oI}$ to $\mathbf{z}^o$ and the vector $(0, \ldots, d_i^o, \ldots, 0)$) or equivalently to a point on the unit circle. A small angle can be interpreted as a high correlation between the inefficiency in one dimension, $\mathbf{d}_i^o$, and the overall inefficiency vector, $\mathbf{d}^o$, i.e. a high inefficiency contribution from dimension i. For efficient units, i.e. with $\|\mathbf{d}^o\| = 0$, the angle is undefined. Note also that the *magnitude* of inefficiencies is the ancillary complement to the angles.

2.2. Operationalization

In order to implement the general definitions above, the technology can be estimated by the use of DEA as shown below, here under the assumption of constant returns to scale.

The coordinates of the MEA ideal reference point, $\mathbf{z}^{oI}$, for $\mathbf{z}^o = (-\mathbf{x}^o, \mathbf{y}^o)$, considering individual improvements on each input and output, are found by solving Problem 1 below for each $i \in \{1, \ldots, m+s\}$:

$$
\begin{aligned}
z_i^{oI} &= \max \delta_i \\
s.t \quad & \\
& \sum_{j=1}^{n} \lambda^j z_i^j \geq \delta_i \\
& \sum_{j=1}^{n} \lambda^j z_{-i}^j \geq z_{-i}^o \qquad -i = 1, \ldots, i-1, i+1, \ldots, m+s (= p) \\
& \lambda^j \geq 0 \qquad\qquad\qquad\qquad j = 1, \ldots, n.
\end{aligned}
$$

The benchmark selection of $\mathbf{z}^o = (-\mathbf{x}^o, \mathbf{y}^o)$ on the efficient frontier in the direction of the ideal point is subsequently found by solving

[1] The use of throughputs, instead of separate input- and output vectors, simplifies the notation and definitions in the following.