

Introducing and Modeling Inefficiency Contributions

Asmild, Mette; Kronborg, Dorte; Matthews, Kent

Document Version

Accepted author manuscript

Published in:

European Journal of Operational Research

DOI:

[10.1016/j.ejor.2015.07.060](https://doi.org/10.1016/j.ejor.2015.07.060)

Publication date:

2016

License

Unspecified

Citation for published version (APA):

Asmild, M., Kronborg, D., & Matthews, K. (2016). Introducing and Modeling Inefficiency Contributions. *European Journal of Operational Research*, 248(2), 725-730. <https://doi.org/10.1016/j.ejor.2015.07.060>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025



Introducing and modelling inefficiency contributions

Mette Asmild, Dorte Kronborg, and Kent Matthews

Journal article (Post print version)

This article was originally published in *European Journal of Operational Research*,
Volume 248, Issue 2, Pages 725-730.

First published online: 6 August 2015.

DOI: [10.1016/j.ejor.2015.07.060](https://doi.org/10.1016/j.ejor.2015.07.060)

Uploaded to Research@CBS: September 2015

Available at: <http://research.cbs.dk/da/publications/introducing-and-modeling-inefficiency-contributions%289e01fb52-c0b3-4c93-a125-6ac5c4e7ab3f%29.html>

© 2015. This manuscript version is made available under the CC-BY-NC-ND 4.0
license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Introducing and modelling inefficiency contributions

Mette Asmild

Department of Food and Resource Economics
University of Copenhagen,

Dorte Kronborg

Center for Statistics, Department of Finance
Copenhagen Business School,

Kent Matthews

Cardiff Business School
University of Cardiff

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.

Keywords: Data Envelopment Analysis (DEA), Productivity and competitiveness, Multi-directional Efficiency Analysis (MEA), directional data, Chinese Banks.

Correspondence: Mette Asmild, IFRO, University of Copenhagen, Rolighedsvej 25, 1958 Frederiksberg C., Denmark. Phone +45 35336886, E-mail: meas@ifro.ku.dk.

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 and 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 and Hougaard, 2013).

The use of Multi-directional Efficiency Analysis (MEA) (see Bogetoft and Hougaard, 1999; Asmild et al., 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 government ownership of banks and weak economic development and low bank efficiency (La Porta et al., 2002 and Barth et al., 2001), one might hypothesize that state-owned banks in China have relatively more inefficiency on non-performing loans than joint-stock banks (Asmild and Matthews, 2012). For further evidence see also Cornett et al. (2010). Similarly, the evidence from the public choice literature suggest that state-owned enterprises, 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 $z_i^{oI} = \max \{z'_i \in \mathbb{R} \mid (z'_i, \mathbf{z}_{-i}^o) \in \mathcal{P}\}$, with $-i = \{1, \dots, i-1, i+1, \dots, m+s\}$. 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 $\mathbf{z}^{oB} = \mathbf{z}^o + \beta^*(\mathbf{z}^{oI} - \mathbf{z}^o)$.

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

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 $\mathbf{d}^o = (\mathbf{z}^{oI} - \mathbf{z}^o) \in \mathbb{R}_+^{p=m+s}$ 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 $\mathbf{d}^o = (d_1^o, \dots, d_p^o)$ 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 $\|\mathbf{d}^o\| = \sqrt{(d_1^o)^2 + \dots + (d_p^o)^2}$ (the length of the diagonal in the box) and $p - 1$ angular coordinates. Let \mathbf{d}_i^o be the vector $(0, \dots, d_i^o, \dots, 0)$ of length p . The relation $\theta_i^o = \arccos(\mathbf{d}_i^o \cdot \mathbf{d}^{oT} / \|\mathbf{d}_i^o\| \|\mathbf{d}^o\|) = \arccos(d_i^o / \|\mathbf{d}^o\|)$ ($\theta_i^o \in [0, \frac{\pi}{2}]$) 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 $(0, \dots, d_i^o, \dots, 0)$) 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 $i \in \{1, \dots, m + s\}$:

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

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 Problem 2 below:

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

and subsequently the benchmark is given as $\mathbf{z}^{oB} = \mathbf{z}^o + \beta^o(\mathbf{z}^{oI} - \mathbf{z}^o)$. Dimension-

specific absolute and relative inefficiencies for $\mathbf{z}^o = (-\mathbf{x}^o, \mathbf{y}^o)$ can now be calculated as $AI_i = z_i^{oB} - z_i^o$ and $RI_i = \frac{z_i^{oB} - z_i^o}{z_i^o}$ respectively.

The main components of the suggested method are illustrated in a 1-input,1-output case in Figure 1 below.

INSERT FIGURE 1 ABOUT HERE

In Figure 1 we see, for the observation \mathbf{z}^o , the corresponding ideal reference point, \mathbf{z}^{oI} , and benchmark selection on the estimated production frontier, \mathbf{z}^{oB} . The inefficiency contribution from the input dimension x is determined by the cosine of the angle θ_x^o . Note that the angle θ_x^o is independent of whether we consider the vector from \mathbf{z}^o to the ideal reference point, \mathbf{z}^{oI} , or to the corresponding benchmark, \mathbf{z}^{oB} .

Note that the general idea of the approach is not limited to considering the inefficiency contributions from one dimension at a time. In the $(p-1)$ dimensional unit sphere (S^{p-1}) , we can express the unit vector by $(p-1)$ angular coordinates e.g. for $p=3$ the spherical coordinates are $(\cos\theta, \sin\theta\cos\varphi, \sin\theta\sin\varphi)$ and the relation to the Cartesian coordinates is that the spherical coordinates are inverse trigonometric functions of relative lengths of sides and vectors. Hereby it is possible to consider combinations of angles, which may be useful for examining hypotheses about more complex patterns within the inefficiencies.

In order to formally analyze inefficiency contributions, e.g. comparing the inefficiency contributions from a given dimension, between different subgroups of production plans, we in the following propose a statistical distribution for the angles (directions).

2.3 Model and method for analysis of directions

Circular distributions have widespread applications in modelling angular data, e.g. in medical, biological or meteorological contexts (Batschelet, 1981; Mardia and Jupp, 2000). The most used distribution is the von Mises-Fisher distribution (Mardia, 1975; Mardia and Jupp, 2000), and its generalizations. The von Mises-Fisher distribution on the $(p - 1)$ dimensional sphere in \mathbb{R}^p , $M_p(\boldsymbol{\mu}, \boldsymbol{\kappa})$, is parameterized with the mean direction $\boldsymbol{\mu}$ ($\|\boldsymbol{\mu}\| = 1$) and the concentration parameter $\boldsymbol{\kappa} \geq 0$. The von Mises-Fisher probability density function for a unit vector \mathbf{x} (i.e. $\mathbf{x} \in \mathbb{R}^p$ and $\|\mathbf{x}\| = 1$) is given by

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\kappa}) = \frac{\boldsymbol{\kappa}^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\boldsymbol{\kappa})} e^{\boldsymbol{\kappa} \boldsymbol{\mu}^T \mathbf{x}},$$

where $I_\nu(\cdot)$ is the modified Bessel function of first kind and order ν .

The von Mises-Fisher distribution can be seen as a spherical analogue to the normal distribution. The concentration parameter $\boldsymbol{\kappa} \geq 0$ measures how concentrated the unit vectors are about the mean direction $\boldsymbol{\mu}$; the larger the more concentrated. For $\boldsymbol{\kappa} = 0$ the density is the uniform density on the sphere.

When $p = 2$ the density reduces to the von Mises distribution (also referred to as the circular normal distribution) for an angle $\theta \in [0, 2\pi)$ with density given by

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}.$$

When κ is large the distribution can be approximated by a normal distribution $\kappa^{\frac{1}{2}}(\theta - \mu) \sim N(0, 1)$ and small κ corresponds to a uniform distribution on the circle.

The appropriate distribution for inefficiency contributions as defined above is a trun-

cated distribution with support restricted to the part of the circle circumference located in the first quadrant. For this reason the von Mises density function above is normalized, and the truncated von Mises distribution has density

$$g(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)} / \int_0^{\pi/2} \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)} dx, \quad (1)$$

where $\theta \in [0, \frac{\pi}{2}]$.

Due to the relatively few observations in the empirical illustration below, the statistical analyses of the inefficiency contributions in this paper are restricted to analyses of each dimension separately, and the distribution used is (1) above or the corresponding truncated normal distribution when appropriate, i.e. when the κ -values are large. The parameters are estimated by maximum likelihood and 95 percent confidence intervals are based on profile likelihood functions. Inference concerning the parameters are carried out as likelihood ratio test, and the resulting -2log LR test statistics are evaluated in the appropriate asymptotic χ^2 -distributions.

3 Empirical illustration

To illustrate the method we use a data set on Chinese banks which is a panel data set consisting of 5 State Owned Commercial Banks (SOCB), 9 Joined Stock Commercial Banks (JSCB) and 13 City Commercial Banks (CCB) each measured over 5 years (2006-2010) for a total of 135 observations. The data is sourced from Bureau van Dijk Bankscope. Efficiency is measured using a simple bank production model with labour and fixed assets as inputs, and net interest revenue and non-interest revenue as outputs. Such a specification closely resembles the neo-classical production function that translates stocks of factor inputs into flows of output, measured by revenue. However, there is no universally accepted

specification of inputs and outputs in bank efficiency analysis. The literature is divided between the intermediation approach of Sealey and Lindley (1977) who use fixed assets, labour, and deposits as inputs and earning assets as outputs, and the production approach identified by Berger and Humphrey (1997), where outputs are the various services of the bank. Largely because of the ease of measurement, the intermediation approach has dominated the literature, but here also there is division on whether deposits should be an input or an output. In a variation of the intermediation approach, Drake and Hall (2003) point to the increasing use of non-interest earnings as an output in bank efficiency studies, recognising the role of off-balance sheet activity in income generation. More recently, Fethi and Pasiouras (2010), in their survey of bank efficiency studies, highlight an increasing number of studies that employ a further variation that has revenue components as outputs. This is the approach taken in this paper. To get sufficient sample size we have pooled observation across the study period. Note that pooling positive correlated observations can cause too low estimates of the concentration parameters and probably less reliable test. However, in our example only neighbour (in time) observations tend to be positively correlated (though not significantly) whereas observations separated by more than one year can be assumed to be independent. Descriptive statistics of the model variables are provided in Table 1 below. It is here worth noting that the input and output values are highest for the SOCBs, and lowest for the CCBs.

Type	n	Mean (St.Dev)			
		Labour	Fixed Assets	Non-IR	Net IR
SOCB	25	291040 (130135)	58580 (22301)	27645 (15689)	143976 (59734)
JSCB	45	18338 (8874)	5191 (2366)	2251 (2092)	20526 (10122)
CCB	65	2525 (1758)	805 (704)	129 (165)	2396 (2712)

Table 1. Descriptive statistics of model variables (all measured in RMB 000s).

3.1 Preliminary analysis

Based on the estimated frontier of the production possibility set, the standard MEA inefficiencies are calculated and presented in Table 2 below. The first set of results are the average absolute inefficiencies within each of the three bank types for each of the four variables. The second set are the corresponding average relative inefficiencies and the third set are the angles corresponding to the definition in Section 2. Three observations, two JSCBs and one CCB, are fully efficient, resulting in undefined angles, thus the average angles for the JSCBs in the tables below are based on 43 observations and for the CCBs on 64 observations.

Type	Absolute inefficiencies			
	Labour	Fixed Assets	Non-IR	Net IR
SOCB	69539	12711	12162	117718
JSCB	1691	766	598	3414
CCB	516	201	164	854

Type	Relative inefficiencies			
	Labour	Fixed Assets	Non-IR	Net IR
SOCB	0.2229	0.2140	0.6130	0.8603
JSCB	0.1077	0.1493	0.6240	0.2033
CCB	0.2398	0.2692	4.1540	0.6688

Type	Angles (in radians)			
	Labour	Fixed Assets	Non-IR	Net IR
SOCB	1.077	1.467	1.487	0.520
JSCB	1.087	1.349	1.397	0.617
CCB	1.050	1.373	1.404	0.626

Table 2. Average absolute inefficiencies, relative inefficiencies, and angles for each bank type.

Considering the absolute inefficiencies reveals that they, for all four variables, are much larger for the SOCBs than for the other bank types. This does not, in fact, provide much

information about the pattern within the inefficiencies because the absolute inefficiencies are affected by both the level of inefficiency as well as the scale of the observations. Regarding the latter, Table 1 above already revealed that the SOCBs generally operate on a much larger scale than the other banks and thus also are likely to have higher absolute inefficiencies.

One way to control for the scale of the observations, is to consider the relative inefficiencies, but note that these values are still affected by the levels of inefficiency. So the fact that the JSCBs generally have the smallest inefficiencies on all variables (except on Non-IR where they are still quite similar to the SOCBs), could be because these banks are generally less inefficient, i.e. located closer to the frontier. In fact, the average input oriented DEA efficiencies in the three groups are 0.54, 0.75 and 0.47 respectively.

As the angles do not depend on the levels of the inefficiencies, considering these enables us to specifically examine the patterns within the inefficiencies. Here we observe, that there seems to be only little difference between the bank types on the inefficiency contribution from labour, whereas especially the SOCBs appear different on the other three variables: the SOCBs have larger angles on Non-IR and on fixed assets, which implies relative less inefficiency on these variables, but have smaller angles, so (relatively) more inefficiency contribution from Net IR than the other bank types.

3.2 Results

The estimated parameters from the truncated von Mises distributions (1) are shown in Table 3 below for each type of bank.

Type	N		Labour	Fixed Assets	Non-IR	Net IR
SOCB	25	$\hat{\mu}$	1.077	1.472	1.487	0.519
		CI	(1.028, 1.125)	(1.448, 1.514)*	(1.478, 1.495)*	(0.4725, 0.5666)
		$\hat{\kappa}$	71.06	321.1	2366	75.44
		CI	(38.63, 118.22)	(125.0, 590.6)*	(1278, 3932)*	(41.03, 125.21)
JSCB	43	$\hat{\mu}$	1.106	1.528	1.401	0.609
		CI	(1.029, 1.207)	(1.367, 1.571)	(1.376, 1.431)	(0.533, 0.682)
		$\hat{\kappa}$	17.44	15.53	156.77	18.66
		CI	(9.448, 27.62)	(8.815, 34.40)	(87.72, 244.64)	(11.17, 28.20)
CCB	64	$\hat{\mu}$	1.051	1.528	1.513	0.625
		CI	(1.017, 1.085)	(1.404, 1.570)	(1.449, 1.583)	(0.588, 0.663)
		$\hat{\kappa}$	54.21	19.89	31.82	44.21
		CI	(37.47, 73.18)	(12.06, 39.17)	(17.52, 61.16)	(30.64, 61.27)

Calculations based on the truncated normal distribution are marked with *

Table 3. Maximumlikelihood estimates for direction parameter, μ , and concentration parameter, κ , in the truncated von Mises distribution (1), together with 95 percent confidence intervals based on the profile likelihood.

First, it is noted that the concentration parameters seems to differ considerably. Table 3 reveals that it is the SOCBs that have the largest concentration, i.e. the least variance, in all dimensions which is to be expected, since state owned banks in China were created in

1979 from the commercial arm of the Peoples Bank of China, which had a monopoly on all banking activities. However, despite modernization, listing and reforms, the policy oriented activity of the state-owned banks, its branches and organizational structure remains much the same as part of its inheritance (Chen et al., 2005).

Dimension	Bank types compared	Direction μ		Concentration κ	
		χ^2	p	χ^2	p
Labour	SOCB/JSCB/CCB	1.976	0.3724	20.15	0.0000
	SOCB/JSCB	0.3990	0.5276	11.89	0.0006
	SOCB/CCB	0.7564	0.38447	0.6452	0.4218
	JSCB/CCB	1.635	0.2010	14.60	0.0001
Fixed Assets	SOCB/JSCB/CCB	0.5821	0.7474	16.206	0.0000
	SOCB/JSCB	0.1952	0.6586	17.43	0.0000
	SOCB/CCB	0.3994	0.5274	15.75	0.0000
	JSCB/CCB	0.000	1.000	0.1370	0.7112
Non-IR *	SOCB/JSCB/CCB	13.48	0.0012	55.10	0.0000
	SOCB/JSCB	13.21	0.0014	31.69	0.0000
	SOCB/CCB	0.3377	0.8446	44.86	0.0000
	JSCB/CCB	6.553	0.0377	112.149	0.0005
Net IR	SOCB/JSCB/CCB	11.525	0.0031	16.941	0.0002
	SOCB/JSCB	3.997	0.0456	12.52	0.0004
	SOCB/CCB	11.271	0.0009	2.372	0.1235
	JSCB/CCB	0.1413	0.7069	9.4212	0.0021

Calculations based on the truncated normal distribution are marked with *

Table 4. Comparison of direction and concentration parameters in the truncated von Mises distribution or truncated normal distribution (*). Highly significant results

($p < 0.01$) are in bold.

The differences in the concentration parameter are confirmed by likelihood ratio tests, the results of which are shown in Table 4. The overall conclusion is that the concentration parameters are significantly larger for SOCBs than for the JSCBs in all dimensions and also significantly larger than those for the CCBs for fixed assets and Non-IR. This means that the distributions of directions are generally most concentrated for SOCBs. Furthermore, the concentrations for the CCBs are significantly different from those for the JSCBs for the dimensions Labour, Non-IR and Net IR.

Regarding the direction parameter μ , the results in Table 4 show, that this parameter for SOCBs in the Non-IR dimension is highly significantly different from that for the JSCBs, whereas neither of the two are significantly different than that for the CCBs. For Net IR this pattern is (almost) reversed, as there are significant difference between directions for SOCBs and CCBs, whereas these are not clearly different from the JSCBs.

To illustrate the results, the truncated von Mises distributions corresponding to the parameters from Table 3 are shown in Figure 2 below. Again it is seen that the SOCBs are the most concentrated in all dimensions but also that the SOCBs and the JSCBs are more concentrated on Non-IR than in the other dimensions.

Furthermore the smallest angles, i.e. largest inefficiency contributions ($\hat{\mu}$), are for Net IR for all three bank types. It can also be noticed that the SOCBs have the larger angle on Non-IR compared to the JSCBs meaning that these banks, tend to have smaller inefficiency contributions from Non-IR, and similarly larger inefficiency contributions on Net IR for SOCBs than JSCBs. SOCBs have historically had low net interest earnings because of its inherited non-performing loans history and the continued use of the state-owned banks to support politically sponsored projects as part of the socialist plan.

INSERT FIGURE 2 ABOUT HERE

Besides the clear patterns in the results discussed above, Table 4 also indicates some tendency towards differences on other parameters, besides those mentioned above. However, due to the limited number of observations, the calculated p-values have to be interpreted with caution, and therefore we here only emphasize the highly significant results.

4 Discussion

In this paper we argue that the understanding of efficiency can be enhanced by looking at patterns within, and not just levels of, the inefficiencies as this enables analysis of different sets of hypothesis, e.g. relating to which variables contribute most to the inefficiency.

Multi-directional Efficiency Analysis (MEA) provides a starting point for analyzing inefficiency patterns. Building on the MEA approach we here first suggest to consider the directions (which can be interpreted as inefficiency contributions) rather than the absolute or relative inefficiencies on the different variables. This enables us to control for the levels of inefficiency and hereby specifically examine the patterns. Secondly, we have proposed a statistical model for analyzing these directions, which has been used here for hypothesis tests relating to differences between subgroups. This has enabled us to identify significant differences between the state owned banks in China, vis-a-vis other bank types, in terms of how much individual input- and output variables contribute to the overall inefficiency. The lesson from this exercise is that the political constraints placed on the directed lending of the SOCBs to loss-making enterprises will significantly contribute to the inefficiency on Net IR. But the influence of minority holding of SOCBs by foreign banks will result in significantly reduced inefficiency in Non-IR as suggested in Berger et al. (2009).

Note that we in the empirical illustration, in order to get sufficient sample size, have pooled observations across 5 years, artificially boosting the sample size. Since the observations for a given bank in different years are positively correlated, this implies that the results from the tests here have to be interpreted with caution. Furthermore, the use of a pooled frontier may be problematic if there are substantial technical changes during the study period as this challenges the homogeneity assumption underlying DEA. Such heterogeneity affects the relevance and appropriateness of identified benchmarks in DEA/MEA in general, and therefore also the ideal points and resulting angles considered here.

In terms of our empirical case of Chinese Banks, since many of the results are actually highly significant, we are still able to identify some clear patterns in spite of the potential problems above. The findings confirm what is generally known about the Chinese banking system. The SOCBs are politically constrained to lend to uneconomic projects as part of the government socialist plan. They have in the past been constrained to maintain branches in uneconomic regions and to employ retirees from the army and absorb party officials in the workforce as a political imperative (see Matthews, 2013). Divesting themselves from these positions takes time but it is clear from where the contributions to inefficiency arise. For future applications, having enough observations to facilitate analysis within each time period separately, would circumvent the problems above but also enable a specific investigation of changes in the inefficiency contributions over time.

Whilst we in this paper have considered MEA inefficiencies, it should be noted that the proposed method for comparing inefficiency contributions is equally applicable to other related techniques, for example the Range Directional Measure (RDM) of Portela, Thanassoulis and Simpson (2004), which also estimates non-radial inefficiencies, but using a different ideal point from that of MEA.

Besides using the directional models to test differences in the distributions of the inefficiency contributions between subgroups, like in the application to Chinese banks above, one might also use (variations of) the approach outlined here to compare the inefficiency contributions from different variables. This could be relevant for e.g. analyzing bargaining power (in terms of identifying which staff groups are allowed more slack), revealed preferences (in terms of which variables it is desirable to have slack on) etc. c.f. the notion of rational inefficiency (Bogetoft and Hougaard, 2003).

It is well-known (see e.g. Banker, 1993) that the convex envelopment of a set of observed production plans is an inner approximation, and thus a biased estimate, of the true production frontier (the boundary of \mathcal{P}) and this, in turn, results in downward biased estimates of the distances to the frontier. Note, however, that our use of angles as measures of inefficiency contributions to a large extent circumvents the problems associated with using an estimated frontier for standard efficiency analysis. Unlike the traditional efficiency scores from DEA, the angles (inefficiency contributions) are not biased, as long as the shape of the estimated frontier is the same as that of the true frontier, i.e. with the same slopes.

For extreme input and/or output mixes, the corresponding facets of the frontier may be poorly estimated, as they are based on few observations, and therefore not necessarily close (in shape) to the true unknown frontier; specifically the axis-parallel extensions of the frontier (or non fully dimensional facets) are potentially problematic. However, the larger the sample size, the better the estimated frontier approximates the true unknown production frontier. Furthermore, even though the non fully dimensional facets are defined by few efficient observations, also inefficient observations located close to these facets support their estimation as approximations of the true frontier. If there are few observations with such extreme input and/or output mixes, making the non fully dimensional facets poor

estimates of the true frontier, then the angles for these observations may exaggerate the differences between subgroups. However, since in those cases only few units contribute to the problem, then the overall comparisons of the subgroups may still not be substantially affected.

Finally, note that applying the suggested approach to directions relative to input-oriented DEA benchmarks (as opposed to MEA benchmarks) means testing for differences in the input mix between the subgroups (and, of course, similarly for output oriented DEA).

References

- Asmild, M., Bogetoft, P. and Hougaard, J.L. (2013), Rationalising Inefficiency: A study of Canadian bank branches, *OMEGA*, 41(1), 80-87.
- Asmild, M., Hougaard, J.L., Kronborg, D. and Kvist, H.K. (2003), Measuring inefficiency via potential improvements, *Journal of Productivity Analysis*, 19, 1, pp. 59-76.
- Asmild, M. and Matthews, K. (2012), Multi-directional Efficiency Analysis of Efficiency Patterns in Chinese Banks 1997-2008, *European Journal of Operational Research*, 219, 434-441.
- Banker, R.D. (1993), Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation, *Management Science*, 39(1), 1265-1273.
- Barth, J.R., Caprio, G.Jr. and Levine, R. (2001), Banking systems around the globe: Do regulations and ownership affect performance and stability? in Mishkin, F.S. (Ed) Prudential Supervision: What Works and What Doesn't, **Univ of Chicago Press**, 31-88.
- Batschelet, E. (1981), Circular Statistics in Biology, **Academic Press**, London.
- Berger A.N., Hasan I. and Zhou, M. (2009), Bank ownership and efficiency in China: What will happen in the world's largest nation? *Journal of Banking and Finance*, 33, 113-130.
- Berger A.N. and Humphrey, D.B. (1997), Efficiency of financial institutions: International survey and directions for future research, *European Journal of Operational Research*, 98, 175-212.
- Bogetoft, P. and Hougaard, J.L. (2003), Rational inefficiencies, *Journal of Productivity Analysis*, 20, 243-271.
- Bogetoft, P. and Hougaard, J.L. (1999), Efficiency evaluations based on potential (non-proportional) improvements, *Journal of Productivity Analysis*, 12, 233-247.
- Charnes, A., Cooper, W.W. and Rhodes, E. (1978), Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research* 2, 429-444.

- Chen, X., Skully, M. and Brown, K. (2005), Banking Efficiency in China: Applications of DEA to pre-and post-Deregulation eras: 1993-2000, *China Economic Review*, 16, 229-245.
- Cornett, M.M., Guo, L., Khaksari, S. and Tehranian, H. (2010), The impact of state ownership on performance differences in privately owned versus stat-owned banks: An international comparison, *Journal of Financial Intermediation*, 19, 74-94.
- Drake, L. and Hall, M.J.B. (2003), Efficiency in Japanese banking: an empirical analysis, *Journal of Banking and Finance*, 27, 891-917.
- Fethi, M.D. and Pasiouras, F. (2010), Assessing Bank Efficiency and Performance with Operational Research and Artificial Intelligence Techniques: A Survey, *European Journal of Operational Research*, 204, 189-198.
- La Porta, R., Lopez-de-Silanes, F. and Shleifer, A.A. (2002), Government ownership of banks, *Journal of Finance*, 57, 265-302.
- Mardia, K.V. (1975), Statistics of Directional Data, *Journal of the Royal Statistical Society, Series B (Methodological)*, 37(3), 349-393.
- Mardia, K.V. and Jupp, P.E. (2000), **Directional Statistics**, Wiley.
- Matthews, K. (2013), Risk Management and Managerial Efficiency in Chinese Banks: A Network DEA Framework, *Omega: The International Journal of Mangement Science*, 41, 207-215.
- Megginson, W.L. (2005), The Economics of Bank Privatization, *Journal of Banking and Finance*, 29, 1931-1980.
- Portela, M.C.A.S., Thanassoulis, E. and Simpson, G. (2004), Negative data in DEA: A directional distance approach applied to bank branches, *Journal of the Operational Research Society*, 55, 1111-1121.
- Sealey, C.W. and Lindley, J.T. (1977), Inputs, outputs, and theory of production cost at depository financial institutions, *Journal of Finance*, 32, 1251-1266.

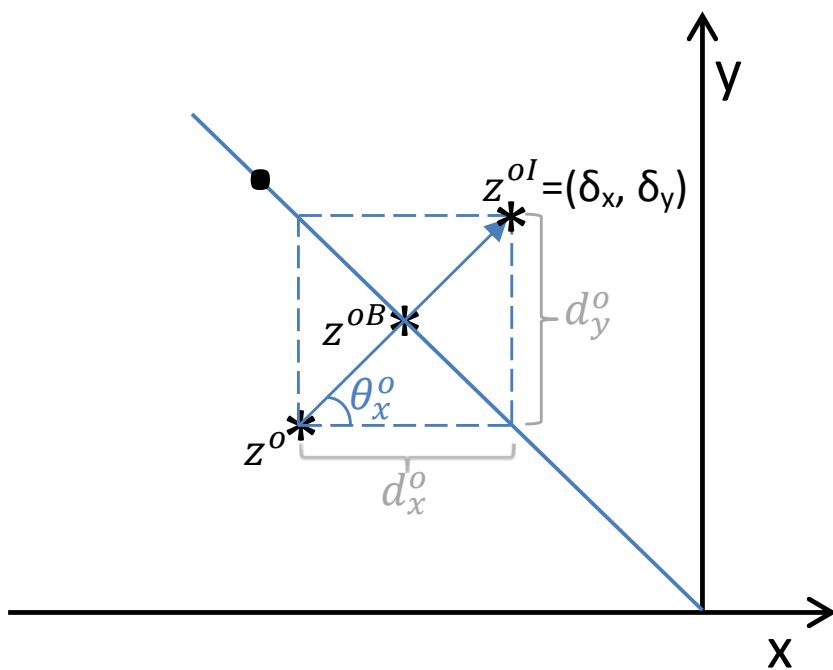


Figure 1: Illustration of the basic elements of the approach.

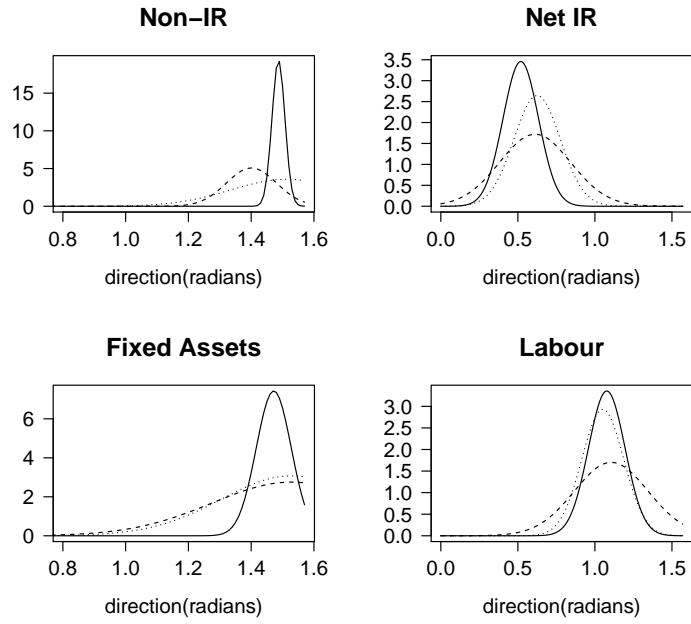


Figure 2: Estimated distributions for directions for each dimension and bank type: SOCB (solid line), JSCB (dashed line), and CCB (dotted line).