Citation for published version

Shen, Wanfang and Yang, Guoliang and Zhou, Zhongbao and Liu, Wenbin (2018) DEA models with Russell measures. Annals of Operations Research . ISSN 0254-5330.

DOI

https://doi.org/10.1007/s10479-018-2867-2

DEA Models with Russell Measures

Wanfang Shen[a], Guoliang Yang[b], Zhongbao Zhou[c,*], Wenbin Liu[d]

a. School of Mathematic and Quantitative Economics, Shandong University of Finance and Economics, Jinan, 250014, China
b. Institute of Science and Development, Chinese Academy of Sciences, Beijing, 100864, China
c. School of Business Administration, Hunan University, Changsha, 410082, China
d. KBS, University of Kent, CT2 7NZ, UK

Abstract: In real applications, data envelopment analysis (DEA) models with Russell measures are widely used although their theoretical studies are scattered over the literature. They often have seemingly similar structures but play very different roles in performance evaluation. In this work, we systematically examine some of the models from the viewpoint of preferences used in their Production Possibility Sets (PPS). We identify their key differences through the convexity and free-disposability of their PPS. We believe that this study will provide guidelines for the correct use of these models. Two empirical cases are used to compare their differences.

Keywords: DEA; Russell measures; Preference; Production possibility sets; Free disposability

1. Introduction

Data envelopment analysis (DEA) is a systematic approach for analyzing the performance of organizations and operational processes, which was first proposed by Charnes, et al. (1978), based on economic theory and linear programming. The DEA models can facilitate comprehensive measurement using input/output data to evaluate the relative efficiency of decision making units (DMUs) without a prior knowledge of input/output functions and weights. Now there are numerous theoretical and empirical researches into the DEA method, which has been extended to many areas, including private sectors and public sectors. Its theories and applications can be found in various books and surveys like Coelli, et al. (2005), Cook, et al. (2007), Cooper, et al. (2004),

* Corresponding author, Z.B.Zhou@hnu.edu.cn, Z.B.Zhou@163.com.

Cooper, et al. (2006), Cooper, et al. (2007), Liu, et al. (2015b), Ray (2004), Zhu (2009), and Zhu, et al. (2007).

In the original DEA models, the radial measures were used. Later on the Russell measures were found to be useful, see Fare and Lovell (1978) for an early study of efficiency measures, and remark in Bogetoft and Hougaard (1999),Levkoff et al. (2012), Russell, et al. (2009) for some studies of the Russell measures. For the DEA models with the Russell measrues, see Aparicio and Pastor (2013), Cooper, et al. (2007), Mirsalehy et al (2014) , Pastor, et al. (1999), Tone (2001), Zhu(1996), Zhu(1998), and Zhou, et al.(2013), which can be used to measure the non-radial part of efficiency.

There are many different DEA models with the Russell measures in the literature. It is often confusing in empirical applications as which should be applied, as although often having seemingly similar structures, they have in fact very different emphases in empirical applications as to be seen later. Therefore, there is a need to examine their characteristics to provide useful advice for the empirical applications. In this work, we will systematically study them from the viewpoint of the preferences used in their Production Possibility Sets (PPS) since a DEA model is essentially determined from its PPS and measure. We find that some of the Russell DEA models use the Pareto preference in their PPS as in the classic DEA models with the radial measure so that their emphasis is to measure non-radial part of efficiency in empirical applications. However, we find that the other Russell DEA models use very different preferences in their PPS, and then their emphasis is not to measure the non-radial part of efficiency, but to measure some compensable sums of the relative measures of the input/output components. Thus one uses them whenever one wishes to compare some total sums of input/output componets, rather than their individual components (more details are in Section 6). Therefore, we are able to provide some useful guidance for using (or not using) which DEA models with the Russell measures in empirical studies.

The structure of the paper is as follows: In Section 2, we will introduce some DEA models with Russell measures. In Section 3 we will discuss the convexity and free-disposability of the PPS for general preferences. Then in Section 4 we examine the convexity and free-disposability of the PPS for the DEA models with the Russell measures. In final section, we compare those models in case studies.

2. Some DEA models with Russell measures

Let $(X_j, Y_j)\ (j = 1, \ldots, n)$ be DMUs and for DMU j, let x_{ij} and y_{rj} represent its i-th ($i=1,\ldots m$) input and r-th ($r=1,\ldots s$) output. In this work, we assume that all the components of inputs/outputs are positive. The standard input and output oriented DEA models with the Russell measure read respectively (Cooper, et al. 2007),

$$\begin{aligned}
&\min\ \frac{1}{m}\sum_{i=1}^{m}\theta_i \\
&\text{s.t.} \\
&\sum_{j=1}^{n}\lambda_j x_{ij} \le \theta_i x_{i0},\ \ i = 1,\ldots,m \\
&\sum_{j=1}^{n}\lambda_j y_{rj} - s_r^{+} = y_{r0},\ \ r = 1,\ldots,s \qquad (2.1) \\
&\lambda_j \ge 0,\ \ j = 1,\ldots,n \\
&s_r^{+} \ge 0,\ \ r = 1,\ldots,s \\
&0 \le \theta_i \le 1,\ \ i = 1,\ldots,m
\end{aligned}$$

$$\begin{aligned}
&\max\ \frac{1}{s}\sum_{r=1}^{s}\theta_r \\
&\text{s.t.} \\
&\sum_{j=1}^{n}\lambda_j x_{ij} + s_i^{-} = x_{i0},\ \ i = 1,\ldots,m \\
&\sum_{j=1}^{n}\lambda_j y_{rj} \ge \theta_r y_{r0},\ \ r = 1,\ldots,s \qquad (2.2) \\
&\lambda_j \ge 0,\ \ j = 1,\ldots,n \\
&s_i^{-} \ge 0,\ \ i = 1,\ldots,m \\
&\theta_r \ge 1,\ \ r = 1,\ldots,s
\end{aligned}$$

As mentioned before, those models are used widely to measure the non-radial part of efficiency for DMUs. It is also possible to introduce weights in the objective functions to express the relative importance of the inputs or outputs. However, these models do not increase their discrimination power in the sense that if a DMU is efficient in a standard DEA model with the radial measure, it also so with the Russel measures. What is more, it is clear that if a DMU is efficient in one of the above models then its component scores θ_i are all the unity even if one of the inputs or outputs does not really perform well relatively. To address these issues, other models with the Russell measures are introduced. The first example reads as follows (see Zhu (1996), assume all the components of inputs/outputs are positive):

$$
\min \ \frac{1}{m}\sum_{i=1}^{m}\theta_i
$$

$$
\begin{cases}
\sum_{j=1}^{n}\lambda_j x_{ij} = \theta_i x_{i0}, \quad i = 1,\Lambda, m \\
\sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ = y_{r0}, \quad r = 1,\Lambda, s \\
\lambda_j \geq 0, \quad j = 1,\Lambda, n \\
s_r^+ \geq 0, \quad r = 1,\Lambda, s \\
\theta_i \geq 0, \quad i = 1,\Lambda, m
\end{cases}
\tag{2.3}
$$

Note that this model loses the usual constraints of the input orientated Russell DEA model: $\{0 \leq \theta_i \leq 1, i = 1,\Lambda, m\}$, so that $\theta_i (i = 1,\Lambda, m)$ may be greater than the unity as shown in the following:

Table 2.1 An illustrative example of Model 2.3

DMUs	DMU_1	DMU_2	DMU_3
Input1	2	13	15
Input2	12	1	3
Output1	3	6	10
Min.Objective	1	1	1
θ_1	1	1	1.4444
θ_2	1	1	0.5556

This is very different from the standard DEA models with the Russell measures, where if the total score is unity then so are all the subscores θ_i. Thus one advantage of this model is that we can further differentiate the performance of efficient DMUs. Zhu (2009) has discussed these in detail.

The second example of the DEA models (Output-orientated) reads as follows (see Zhu(1996)), assume all the components of inputs/outputs are positive):

$$
\max \quad \frac{1}{s}\sum_{r=1}^{s}\theta_r
$$

$$
\begin{cases}
\sum_{j=1}^{n}\lambda_j x_{ij} + s_i^- = x_{i0}, \quad i = 1,\Lambda\ ,m \\
\sum_{j=1}^{n}\lambda_j y_{rj} = \theta_r y_{r0}, \quad r = 1,\Lambda\ ,s \\
\lambda_j \geq 0, \quad j = 1,\Lambda\ ,n \\
s_i^- \geq 0, \quad i = 1,\Lambda\ ,m \\
\theta_r \geq 0, \quad r = 1,\Lambda\ ,s
\end{cases}
\qquad (2.4)
$$

Note that this model loses the usual constraints of the output orientated Russell DEA model: $\theta_r \geq 1$ so that θ_r may be less than the unity, similarly as we showed in the first model. The above two DEA models have been studied in detail in Zhu(1996) and Zhu(2009), where their duals were derived among the other things. These DEA models have been widely used in the literature; see Chen, et al. (2004), Seiford, et al. (1998), Seifort, et al. (2003), and Zhu (2009).

As explained in Zhu (1996), these two models mentioned above look like some minor extensions of the standard Russell DEA models to deal with wider applications. In fact, they are quite different from the standard DEA models 2.1 and 2.2. The first indication can be seen from the following Table 2.2, which represents the computation results for an example of Model 2.4.

It is clear that all the efficiency scores of the DMUs are less than the unity. Thus the DEA scores cannot be directly used as an efficiency measure as they do not satisfy the first axiom given in Fare and Lovell (1978) for a proper efficiency. This seems to indicate that there are some essential differences between the standard DEA models (with either the radial or Russell measures) and those models, which are important to guide proper uses of them and have not been systematically studied until now.

Table 2.2 An illustrative example of Model 2.4

DMUs	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5
Input1	1	1	1	1	1
Output1	1	1	2	6	3
Output2	2	3	3	2	4
Max.Objective	3.5	3.333	1.833	1.25	1.25
θ_1	6	6	3	0.5	2
θ_2	1	0.667	0.667	2	0.5

Much related to Models (2.1-2.4) are the additive models: It was shown in Liu, et al. (2010) that with the translations of $\theta_i = \frac{x_{i0} - s_i^-}{x_{i0}}, i = 1, K, m; \theta_r = \frac{y_{r0} + s_r^+}{y_{r0}}, r = 1, K, s$, the models with Russell measures and with slacks measures are in fact equivalent.

Thus here we introduce two additive models, which will be examined later and shown that they have many advances over Models 2.3-2.4.

$$\min \quad g_1 = 1 - \frac{1}{m}\left(\sum_{i=1}^{m} \frac{s_i^-}{x_{i0}}\right)$$
$$\begin{cases} \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{i0}, \quad i = 1, \Lambda, m \\ \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0}, \quad r = 1, \Lambda, s \\ \lambda_j \geq 0, \quad j = 1, \Lambda, n \\ s_r^+ \geq 0, \quad r = 1, \Lambda, s \\ \sum_{i=1}^{m} s_i^- \geq 0, \\ s_i^-, \text{ free}, \quad i = 1, \Lambda, m \end{cases} \tag{2.5}$$

$$\max \quad g_2 = 1 + \frac{1}{s}\left(\sum_{r=1}^{s} \frac{s_r^+}{y_{r0}}\right)$$
$$\begin{cases} \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{i0}, \quad i = 1, \Lambda, m \\ \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0}, \quad r = 1, \Lambda, s \\ \lambda_j \geq 0, \quad j = 1, \Lambda\ n \\ s_i^- \geq 0, \quad i = 1, \Lambda, m \\ \sum_{r=1}^{s} s_r^+ \geq 0, \\ s_r^+, \text{ free}, \quad r = 1, \Lambda, s \end{cases} \tag{2.6}$$

Let us note that in the input oriented model (2.5), the slacks for the inputs may be negative and so are those for outputs in Model (2.6). Thus they are closely related to Models (2.3-2.4). However, it will be shown later that they are in fact quite different from the viewpoint of the convexity and preferences of their PPS. Furthermore, if we constraint the slacks to be non-negative, they will boil down to Models (2.1-2.2) see

Liu, et al. (2010).

3 Preference and free-disposability in PPS of DEA models

It is well-known that a DEA model can be completely presented by its PPS and its measure as below. Let $(X_j, Y_j)\ (j = 1, \Lambda\ , n)$ be DMUs, and if we use a radial measure with P representing the PPS, the output oriented DEA model with the radial measure reads: $\max_{\theta \geq 1} \theta \{(X_0, \theta Y_0) \in P\}$.

The production possibility set is defined by all (X,Y) where X can produce Y under certain technology. Let us emphasize that the PPS is the key to link DEA models with the economic theory where two key properties for the PPS are the convexity and free-disposability, see Cooper, et al. (2007), Pastor, et al. (1999), Russell, et al. (2005), Tone (2001). Suppose that $(X_j, Y_j)(j = 1, \Lambda\ , n)$ are DMUs and S is the technology set, then the minimum PPS is the convex expansion of virtual combinations of the DMU data set:

$$P = \{(X,Y) : X = \sum_{j=1}^{n} \lambda_j X_j,\ Y = \sum_{j=1}^{n} \lambda_j Y_j,\ \lambda = (\lambda_1, \Lambda\ , \lambda_n) \in S\},$$

where $S = \{\lambda = (\lambda_1, \Lambda\ , \lambda_n) : \lambda_j \geq 0, \sum_{j=1}^{n} \lambda_j = (\geq, \leq) 1\}$ or $S = \{\lambda = (\lambda_1, \Lambda\ , \lambda_n) : \lambda_j \geq 0\}$.

It is clear that P is convex. Then often we have to further expand this PPS by e.g. free-disposability assumption. To this end we need define the so-called preferences or orders $\{\phi\}$ and $\{\pi\}$. Then the so-called strongly free disposability assumption reads (or simply P is free-disposable):

$$\text{If } (X,Y) \in P,\ Z\ \phi\ X, W\ \pi\ Y,\ \ \text{then} (Z,W) \in P\ .$$

Let us emphasize that these orders are not necessarily the Pareto preference – that is, $\{\phi\}$ or $\{\pi\}$ does not necessarily mean the component-wise inequalities as in the preference of Pareto (see below). Of course, in the classic DEA, the Pareto order is used, and $\{\phi\}$ and $\{\pi\}$ are just the standard vector inequalities $"\geq","\leq"$. Then PPS in the classic DEA models can be presented by

$$PS^* = \{(X,Y): X \geq \sum_{j=1}^{n} \lambda_j X_j,\ Y \leq \sum_{j=1}^{n} \lambda_j Y_j,\ \lambda = (\lambda_1, K\ , \lambda_n) \in S\}.$$

Then DEA models can be derived from this PS* as shown above. In the classic DEA framework, where $\{\phi\}$ and $\{\pi\}$ are understood in the preference of Pareto (i.e., inequalities in componets), it is easy to check that PS* is convex and free-disposable. It has been realized however that some preferences other than the Pareto are useful in applications (Liu, et al.(2006)), and indeed they (e.g. lexicographical preference and matrix preference) have been used in e.g. Olympic game, research evaluation (Liu, et al.(2006, 2010), Zhang, et al. (2009)). For the cases where a non-Pareto preference $\{\phi\}/\{\pi\}$ is used, the PPS with the strong free-disposability has the following form:

$$PS^* = \{(X,Y): X\ \phi \sum_{j=1}^{n} \lambda_j X_j,\ Y\ \pi \sum_{j=1}^{n} \lambda_j Y_j,\ \lambda = (\lambda_1, \Lambda\ , \lambda_n) \in S\}.$$

However, now the set PS* may not be convex as to be seen later, and the strong free disposability cannot be described by the nonnegative slacks conditions (as in the case of the Pareto). It will become clear later that the different behaviors of the above models are reflected by convexity and free-disposability of their PPS in certain preferences. In order to fully address the origins of the differences, we will examine briefly some basic facts on preferences below.

3.1 Preferences and properties of PPS

A preference is a relationship defined for some pairs (x, y) on a set X, which can be denoted by $\{\phi\}$ and $\{\pi\}$ to represent “better than”, and “worse than” (see Liu, et al. (1985)), respectively. That is，for all $x, y \in X$，if $x\ \phi\ y$，then “ x is at least as good as y"; if $x \pi y$, then "x is at most as good as y". The definition of preference looks slightly abstract, but essentially it just clarifies the precise meanings for the vague expressions like "better, worse". Clearly one should have some understandings of these meanings before an evaluation is carried out via DEA. The most classic example is the numerical order (preference) for the real numbers like "$5 \geq 3$" and "$4 \leq 6$". Such an order can be generalized to a column or a table of real numbers – like the Pareto preference to be seen below. However unlike the real number preference, a pair (x, y) generally may not have such a relationship under these generalized preferences: many pairs may not be comparable under these preferences.

When there exist no other elements in X, which are better than an element x, it will be considered as "optimal or non-dominant" in X, although this does not really mean that it is better than the others in assigned preference like in the real numbers, since this could only mean there are many elements incomparable with it. In a sense, a standard DEA model is to find "optimal" DMUs in PPS under the Pareto preference, see, Cooper, et al. (2004) and Liu, et al. (2006).

Most of the preferences used in applications are **Reflexive:** For all x in X, $x \phi x$, and **Transitive:** For all x, y, and z in X, if $x \phi y$ and $y \phi z$, then $x \phi z$. These properties usually hold in real-life applications of DEA, and they are defined to make sure the mathematical summaries of the value judgments of DMUs are consistent. Another important property of a preference is **compatibility for linear operation**: that is to say, translations and multiplication by a positive number preserve the preference structure (Schaefer and Wolff (1999)). We say: the preference $"\phi"$ is linearly compatible if :

$$x \phi y, z \phi w \Rightarrow \lambda x + \mu z \phi \lambda y + \mu w, \forall \lambda, \mu \geq 0.$$

There are other properties of preferences associated with the continuity, which relate to the openness or closeness of the PPS, and will not be discussed here.

Example 1: Pareto preference

The Pareto preference is by far the most widely used one in economic and management areas. We will keep using the usual inequality symbols "$\geq$", "$\leq$" for this preference. Let $y_1 = (y_{11}, \ldots, y_{s1})$, $y_2 = (y_{12}, \ldots, y_{s2})$ be two outputs. Then in the Pareto preference, $y_1 \succ y_2$ $(y_1 \prec y_2)$, or y_1 is better than (worse than) y_2, if and only if $y_{r1} \geq y_{r2} (y_{r1} \leq y_{r2})$ for $r = 1, \ldots, s$. Clearly, Pareto preference is linearly compatible and transitive.

Assuming the strong free-disposability in Pareto preference for inputs and outputs, then the standard PPS in DEA theory reads:

$$PS^* = \{(X, Y) : X \geq \sum_{j=1}^{n} \lambda_j X_j, Y \leq \sum_{j=1}^{n} \lambda_j Y_j, \lambda = (\lambda_1, \ldots, \lambda_n) \in S\}$$

It is well-known that PS* is convex and free-disposable. The DEA models can be described by PS* and the measures to be used. Let us note that Models (2.1) and (2.2) can be expressed as:

$$\max \quad \frac{1}{r} \sum_{r=1}^{s} \theta_r : \{(X_0, \Theta Y_0) \in PS^*, \Theta = \mathrm{diag}(\theta_1, \ldots, \theta_s), \theta_i \geq 1\}.$$

while Models (2.3) and (2.4) cannot be so, as ΘY_0 may not represent an improvement of Y_0 under Pareto preference used in PS* unless $\theta_i \geq 1$. Thus different preferences need exploring. For a general preference $\succeq$, let

$$PS = \{(X, Y) : X \succeq \sum_{j=1}^{n} \lambda_j X_j, Y \preceq \sum_{j=1}^{n} \lambda_j Y_j, \lambda = (\lambda_1, \ldots, \lambda_n) \in S\},$$

then a DEA model can also be described by (if a Russell measure adopted):

$$\max \quad \frac{1}{r} \sum_{r=1}^{s} \theta_r : \{(X_0, \Theta Y_0) \in PS, \Theta = \mathrm{diag}(\theta_1, \ldots, \theta_s)\}.$$

or

$$\max \frac{1}{s}\sum_{r=1}^{s}\theta_r$$
$$\begin{cases} \sum_{j=1}^{n}\lambda_j X_j \prec X_0 \\ \sum_{j=1}^{n}\lambda_j Y_j \succ \Theta Y_0 \\ \lambda \in S,\ \Theta = \mathrm{diag}(\theta_1, \ldots, \theta_s). \end{cases} \tag{3.1}$$

However for a general preference, is PS convex and free-disposable? Below we will show that the answer depends on some properties of the preference used in the PS.

Proposition 3.1: If the preference used in PS is linearly compatible then PS is convex. It is free-disposal if the preference is transitive.

The proof is obvious, so it is skipped here. Otherwise the set PS may be non-convex or non-free-disposable, if the preference used is non-transitive or not linearly compatible. Furthermore for the non-transitive preferences, there may be contradictive loops in the sense that there are DMUs $Z_i (i = 1, \ldots, p, p \geq 2)$ which satisfy $Z_1 \succ Z_2 \succ \cdots Z_p \succ Z_1$, where $Z_j \neq Z_k, j \neq k; 1 \leq j, k \leq p$. In this case the meaning of comparison between DMUs is unclear, which should be an essential idea that underpins DEA theory. However, if the preference is not transitive but there are not contradictive loops then the PS is free-disposable in a weak sense:

$$\{\text{If } (X,Y) \in PS : Z \succ X \succ \sum_{j=1}^{n}\lambda_j X_j, W \prec Y \prec \sum_{j=1}^{n}\lambda_j Y_j \Rightarrow (Z,W) \in PS \quad \text{if } (Z \text{ and } \sum_{j=1}^{n}\lambda_j X_j),$$

$$\text{and } (W \text{ and } \sum_{j=1}^{n}\lambda_j Y_j) \text{ are both comparable } \}.$$

It is clear that if PS is free-disposable in this weak sense then the resulting DEA model is the same as assuming the strong free-disposability.

4. PPS of DEA Models with Russell measures

Below we will examine the convexity and free-disposability of the PPS and the preferences used in the Models 2.1-2.6. It is clear that for Models 2.1-2.2, the preference used is the Pareto preference and thus their PPS are convex and free-disposable. Let us then examine whether it would be the same for the Models 2.3-2.6.

4.1 θ-output preference

Now we can identify the preference underlying the outputs of the model (2.4) as follows:

Definition 4.1: (θ-output preference). For the outputs $Y_0, Y_1 \in R_+^s$, $Y_0 \prec Y_1$, if and only if there exist $\theta_1, \ldots, \theta_s > 0$ such that $\begin{pmatrix} y_{11} \\ \vdots \\ y_{s1} \end{pmatrix} = \begin{pmatrix} \theta_1 y_{10} \\ \vdots \\ \theta_s y_{s0} \end{pmatrix}$, and $\frac{1}{s}\sum_{r=1}^{s}\theta_r \geq 1$.

It is clear that Model 2.4 uses this preference for the outputs, and the Pareto preference for the inputs, that is; using these two references, (2.4) can be expressed as (3.1). However, this preference is not transitive. For example, let's assume three DMUs with outputs $\begin{pmatrix} 6 \\ 2 \end{pmatrix}, \begin{pmatrix} 12 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. It is easy to verify that $\begin{pmatrix} 6 \\ 2 \end{pmatrix} \prec \begin{pmatrix} 12 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 12 \\ 1 \end{pmatrix} \prec \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, however $\begin{pmatrix} 6 \\ 2 \end{pmatrix} \prec \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ does not hold. Furthermore there is a contradictive loop in this preference: $\begin{pmatrix} 6 \\ 2 \end{pmatrix} \prec \begin{pmatrix} 12 \\ 1 \end{pmatrix} \prec \begin{pmatrix} 3 \\ 2 \end{pmatrix} \prec \begin{pmatrix} 6 \\ 2 \end{pmatrix}$. This may make comparison in this order senseless.

The preference is not linearly compatible. For example, here we have 4 DMUs $Y_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, Y_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, Y_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, Y_4 = \begin{pmatrix} 2 \\ 5.5 \end{pmatrix}, Y_1 \prec Y_2, Y_3 \prec Y_4$, but $(Y_1 + Y_3) \prec (Y_2 + Y_4)$ does not hold.

Below we further show that due to the above defects, PS set may be non-convex and non-free-disposable. Let $(a, b) \in R_+^2$ and let S(a ,b) include all the vectors that are smaller than it: $S(a,b)=\{Y=(y_1,y_2)\in R_+^2:(y_1,y_2)\ p\ (a,b)\}$, and it follows the definition 4.1

$$S(a,b)=\{Y=(y_1,y_2)\in R_+^2:\frac{a}{y_1}+\frac{b}{y_2}\geq 2\}$$

This is an area surrounded by the hyperbolic line: $(ay_2+by_1=2y_1y_2)$ (with asymptotic lines $(y_1=\frac{a}{2},y_2=\frac{b}{2}))$, and by axis $(y_1=0,y_2=0,)$. Now let the input be the unity and outputs be (1,2) and (2,1), along with their convex combinations: $\{(2-\lambda,1+\lambda):0\leq\lambda\leq 1\}$. Then $PS=\{Y=(y_1,y_2)\in R_+^2:\frac{2-\lambda}{y_1}+\frac{1+\lambda}{y_2}\geq 2,0\leq\lambda\leq 1\}$ shown as follows (shadowed area with doted boundary）:

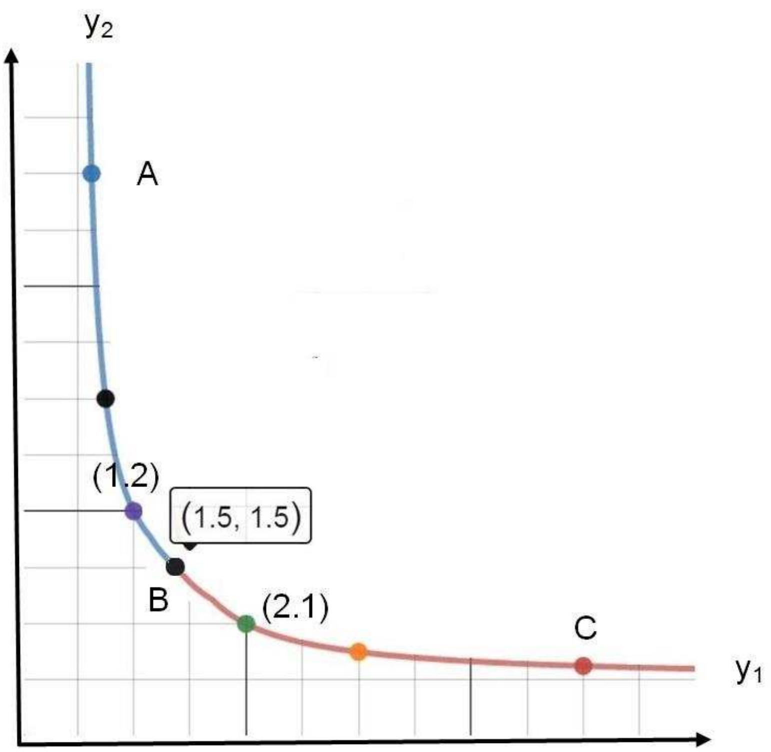

Figure 4.1 Production Possibility Set for θ-output preference

Here the boundary passing through points A and (1,2) is made of the hyperbolic line: $y_2+2y_1=2y_1y_2$ that goes through points B, (1,2) and A; the boundary passing through (2,1) and point C is made of hyperbolic line: $2y_2+y_1=2y_1y_2$ that goes through the points B, (2,1), and C; the boundary passing through points (1,2), B and

(2,1) is made of envelopment of the lines $(2-\lambda)\ y_2 + (1+\lambda)\ y_1 = 2y_1y_2$ that all go through points B, and $\{(2-\lambda, 1+\lambda) : 0 \le \lambda \le 1\}$, the convex combination of points (1,2) and (2,1). It is clear that PS is non-convex. Furthermore PS is also non free-disposable by noting this fact: let（a, b） be a point on the boundary of PS, then usually the set S(a, b) will contain points outside PS，as often the boundary of S(a, b) differs from that of PS. The above discussion holds for the case CRS as well.

Thus the PS of Model (2.4) may be neither free-disposable nor convex, and what is more, the meaning of the comparison underpinning the DEA model seems confusing. Thus it should be avoided to use if possible. This conclusion is supported by the later empirical studies.

Note that if we change the definition into $Y_0 \text{ p } Y_1$ if $\theta_1, \ldots, \theta_s \ge 1$, it is just Pareto preference. Then the model becomes a standard one (2.2).

4.2 θ-input preference

Now we can identify the preference underlying the outputs of the model (2.3) as follows:

Definition 4.2: (θ-input preference). For the inputs $X_0, X_1 \in R_+^m$, $X_0 \ \pi \ X_1$, if and only if there exists $\theta_1, \ldots, \theta_m > 0$ we will have $\begin{pmatrix} x_{10} \\ \vdots \\ x_{m0} \end{pmatrix} = \begin{pmatrix} \theta_1 x_{11} \\ \vdots \\ \theta_m x_{m1} \end{pmatrix}$, and $\frac{1}{m}\sum_{i=1}^{m} \theta_i \le 1.$

It is clear that Model 2.3 uses this preference for the inputs, and the Pareto preference for the outputs. Obviously, the preference defined above is not transitive, for example, $\begin{pmatrix} 3 \\ 4 \end{pmatrix} \pi \begin{pmatrix} 6 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 6 \\ 3 \end{pmatrix} \pi \begin{pmatrix} 14 \\ 2 \end{pmatrix}$, but $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 14 \\ 2 \end{pmatrix}$ are not comparable in this preference. Furthermore $\begin{pmatrix} 3 \\ 4 \end{pmatrix} \pi \begin{pmatrix} 6 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 6 \\ 3 \end{pmatrix} \text{ p } \begin{pmatrix} 14 \\ 3 \end{pmatrix}$, but $\begin{pmatrix} 9 \\ 7 \end{pmatrix} \text{Not p} \begin{pmatrix} 20 \\ 6 \end{pmatrix}$, so this preferences is not linearly compatible either.

However, we can prove that there are no contradictive loops for this preference as stated in the following **Theorem 4.1.**

Theorem 4.1: There are no contradictive loops in θ-input preference. That is to say, in θ-input preference, there exists none of such DMUs with inputs $(X_i, i = 1, \cdots, p; p \geq 2)$ and $X_1 \succ X_2 \succ \cdots \succ X_p \succ X_1$, where $X_j \neq X_k$, $j \neq k$, $1 \leq j, k \leq p$

Proof: We assume that there are p DMUs, whose inputs form a contradictive loop in θ-input preference. Assume that each DMU has m inputs, that is;

$$X_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{m1} \end{pmatrix}, X_2 = \begin{pmatrix} x_{12} \\ \vdots \\ x_{m2} \end{pmatrix}, \cdots, X_p = \begin{pmatrix} x_{1p} \\ \vdots \\ x_{mp} \end{pmatrix} \tag{4.1}$$

Without losing generality, suppose $X_1 \succ X_2 \succ \cdots X_p \succ X_1$, thus we have

$$\begin{cases} \dfrac{x_{11}}{x_{12}} + \dfrac{x_{21}}{x_{22}} + \cdots + \dfrac{x_{m1}}{x_{m2}} \leq m \\ \dfrac{x_{12}}{x_{13}} + \dfrac{x_{22}}{x_{23}} + \cdots + \dfrac{x_{m2}}{x_{m3}} \leq m \\ \vdots \\ \dfrac{x_{1,p-1}}{x_{1,p}} + \dfrac{x_{2,p-1}}{x_{2,p}} + \cdots + \dfrac{x_{m,p-1}}{x_{m,p}} \leq m \\ \dfrac{x_{1,p}}{x_{11}} + \dfrac{x_{2,p}}{x_{21}} + \cdots + \dfrac{x_{m,p}}{x_{m1}} \leq m \end{cases} \tag{4.2}$$

Then we have

$$\frac{x_{11}}{x_{12}} + \frac{x_{21}}{x_{22}} + \cdots + \frac{x_{m1}}{x_{m2}} + \frac{x_{12}}{x_{13}} + \frac{x_{22}}{x_{23}} + \cdots + \frac{x_{m2}}{x_{m3}} + \frac{x_{1,p-1}}{x_{1,p}} + \frac{x_{2,p-1}}{x_{2,p}} + \cdots + \frac{x_{m,p-1}}{x_{m,p}} + \frac{x_{1,p}}{x_{11}} + \frac{x_{2,p}}{x_{21}} + \cdots + \frac{x_{m,p}}{x_{m1}} \leq p*m \tag{4.3}$$

Obviously, (4.3) is equivalent to

$$\left(\frac{x_{11}}{x_{12}}+\frac{x_{12}}{x_{13}}+\cdots+\frac{x_{1,p-1}}{x_{1,p}}+\frac{x_{1,p}}{x_{11}}\right)+\left(\frac{x_{21}}{x_{22}}+\frac{x_{22}}{x_{23}}+\cdots+\frac{x_{2,p-1}}{x_{2,p}}+\frac{x_{2,p}}{x_{21}}\right)+\cdots+\left(\frac{x_{m1}}{x_{m2}}+\frac{x_{m2}}{x_{m3}}+\cdots+\frac{x_{m,p-1}}{x_{m,p}}+\frac{x_{m,p}}{x_{m1}}\right)\le p*m \qquad (4.4)$$

Note that

$$\left(\frac{x_{11}}{x_{12}}+\frac{x_{12}}{x_{13}}+\cdots+\frac{x_{1,p-1}}{x_{1,p}}+\frac{x_{1,p}}{x_{11}}\right)\ge p\cdot\left(\frac{x_{11}}{x_{12}}\cdot\frac{x_{12}}{x_{13}}\cdots\frac{x_{1,p-1}}{x_{1,p}}\cdot\frac{x_{1,p}}{x_{11}}\right)^{\frac{1}{p}}=p\cdot(1)^{\frac{1}{p}}=p,$$

$$\vdots$$

$$\left(\frac{x_{m1}}{x_{m2}}+\frac{x_{m2}}{x_{m3}}+\cdots+\frac{x_{m,p-1}}{x_{m,p}}+\frac{x_{m,p}}{x_{m1}}\right)\ge p, \qquad (4.5)$$

Since the product of all the items inside each pair of brackets is the unity, we have:

$$\left(\frac{x_{11}}{x_{12}}+\frac{x_{12}}{x_{13}}+\cdots+\frac{x_{1,p-1}}{x_{1,p}}+\frac{x_{1,p}}{x_{11}}\right)=p,\cdots,\left(\frac{x_{m1}}{x_{m2}}+\frac{x_{m2}}{x_{m3}}+\cdots+\frac{x_{m,p-1}}{x_{m,p}}+\frac{x_{m,p}}{x_{m1}}\right)=p$$

And further note the well-known fact that the above equalities are true if and only if all the items inside the brackets are the unity.

$$\left(\frac{x_{11}}{x_{12}}=\frac{x_{12}}{x_{13}}=\cdots=\frac{x_{1,p-1}}{x_{1,p}}=\frac{x_{1,p}}{x_{11}}=1\right),\cdots,\left(\frac{x_{m1}}{x_{m2}}=\frac{x_{m2}}{x_{m3}}=\cdots=\frac{x_{m,p-1}}{x_{m,p}}=\frac{x_{m,p}}{x_{m1}}=1\right),$$

Thus we have

$$x_{11}=x_{12}=\cdots=x_{1p},x_{21}=x_{22}=\cdots=x_{2p},\cdots,x_{m1}=x_{m2}=\cdots=x_{mp},$$

So we have $X_1=X_2=\cdots=X_p$, which conflicts the assumption.

Thus PPS of Model (2.3) is free-disposable in the weak sense. Furthermore, we have the following theorem:

Theorem 4.2: There will always be at least one DMU for the **Model 2.3** with the unity being its efficiency score.

Before we prove this theorem, we first need the following two lemmas. Consider the following two linear programming problems,

$$
(P)\begin{cases}
\min \frac{1}{m}\sum_{i=1}^{m}\theta_i \\
s.t. \quad \sum_{j=1}^{n}\lambda_j Y_j \geq Y_0 \\
\sum_{j=1}^{n}\lambda_j X_j = \begin{pmatrix} \theta_1 & & & & \\ & \theta_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & \theta_m \end{pmatrix} X_0 \\
\theta_i \geq 0, i = 1,\ldots,m, \quad \lambda_j \geq 0, j = 1,\ldots,n
\end{cases}
\tag{4.6}
$$

The dual model of (4.6) reads (see Zhu (1996)):

$$
(D)\begin{cases}
\max \quad \sum_{r=1}^{s}\mu_r y_{r0} \\
s.t \quad \sum_{i=1}^{m}\omega_i x_{ij} - \sum_{r=1}^{s}\mu_r y_{rj} \geq 0 \quad , \quad j = 1,\ldots,n \\
\omega_i x_{i0} = \frac{1}{m} \quad , \quad i = 1,\ldots,m \\
\mu_r \geq 0,\ \omega_i \geq 0 \quad , \quad r = 1,\ldots,s;\ i = 1,\ldots,m
\end{cases}
\tag{4.7}
$$

Lemma 4.1: We assume that $\hat{\omega} \geq 0, \hat{\mu} \geq 0, (\hat{\omega}^T, \hat{\mu}^T) \neq 0$ is a solution of Model 4.7. If $\sum_{i=1}^{m}\hat{\omega}_i x_{ij_*} - \sum_{r=1}^{s}\hat{\mu}_r y_{rj_*} = 0$ holds for a certain $j_*(1 \leq j_* \leq n)$, then the efficiency score of DMU_{j*} is the unity.

Proof. Because $\sum_{r=1}^{s}\mu_r y_{r0} \leq \sum_{i=1}^{m}\omega_i x_{i0} = \sum_{i=1}^{m}\frac{1}{m} = 1$, we have $\sum_{r=1}^{s}\mu_r y_{r0} \leq 1$. That is to say that the optimal objective value of **Model 4.7** is no more than 1. Because $\sum_{r=1}^{s}\hat{\mu}_r y_{rj_*} = \sum_{i=1}^{m}\hat{\omega}_i x_{ij_*} = \sum_{i=1}^{m}\frac{1}{m} = 1$, $(\hat{\omega}^T, \hat{\mu}^T)$ is the optimal solution of **Model 4.7**. So, the efficiency score of DMU_{j*} is the unity.

Lemma 4.2: $\exists\bar{\varepsilon}>0$ and $\forall\varepsilon\in(0,\bar{\varepsilon})$, the feasible solutions of **Model 4.8** exist.

$$(D^*)\begin{cases} \max & \sum_{r=1}^{s}\mu_r y_{r0} \\ s.t & \sum_{i=1}^{m}\omega_i x_{ij}-\sum_{r=1}^{s}\mu_r y_{rj}\ge 0 \quad , \quad j=1,\ldots,n \\ & \omega_i x_{i0}=\dfrac{1}{m} \quad , \quad i=1,\ldots,m \\ & \omega_i\ge\varepsilon,\ \mu_r\ge\varepsilon \quad , \quad i=1,\ldots,m;\ r=1,\ldots,s \end{cases} \tag{4.8}$$

Proof. Let

$$\hat{\omega}_i=\frac{1}{mx_{i0}},i=1,\ldots,m$$
$$\hat{\mu}_r=\varepsilon,r=1,\ldots,s$$
$$\varepsilon\in(0,\bar{\varepsilon})$$

where $$\bar{\varepsilon}=\min\{\min_{1\le i\le m}\hat{\omega}_i,\min_{1\le j\le n}\frac{\sum_{i=1}^{m}\hat{\omega}_i x_{ij}}{\sum_{r=1}^{s}y_{rj}},1\}$$

Thus we have

$$\sum_{i=1}^{m}\hat{\omega}_i x_{ij}-\sum_{r=1}^{s}\hat{\mu}_r y_{rj}=\sum_{i=1}^{m}\hat{\omega}_i x_{ij}-\sum_{r=1}^{s}\varepsilon y_{rj}\ge 0, j=1,\ldots,n$$

$$\hat{\omega}_i x_{i0}=\frac{1}{m},i=1,\ldots,m$$

$$\hat{\omega}_i\ge\varepsilon,\hat{\mu}_r\ge\varepsilon$$

Thus we know $\hat{\omega}_\mathrm{i},\hat{\mu}_\mathrm{r}$ is a feasible solution of **Model 4.8**.

Here we begin to prove the Theorem 4.2. Firstly, consider the following **Model**.

$$
(D')\begin{cases} \max & \sum_{r=1}^{s}\mu_r y_{r0} \\ s.t & \sum_{i=1}^{m}\omega_i x_{ij}-\sum_{r=1}^{s}\mu_r y_{rj}\ge 0 \quad , \quad j=1,\ldots,n \\ & \omega_i x_{i0}=\dfrac{1}{m} \quad , \quad i=1,\ldots,m \\ & \omega_i \ge \varepsilon,\ \mu_r\ge\varepsilon \quad , \quad i=1,\ldots,m;\ r=1,\ldots,s \end{cases} \tag{4.9}
$$

From **Lemma 4.2**, we know $\exists\varepsilon>0$ and there are feasible solutions in Model 4.9. Because $\sum_{r=1}^{s}\mu_r y_{r0}\le\sum_{i=1}^{m}\omega_i x_{i0}=1$, we know that the optimal solution of Model 4.9 exists. We let $\omega_i^{\;0},\mu_r^{\;0},i=1,\ldots,m,r=1,,\ldots,s$.

(i) If $\exists j_*(1\le j_*\le n)$ satisfies $\sum_{i=1}^{m}\omega_i x_{ij_*}-\sum_{r=1}^{s}\mu_r y_{rj_*}=0$, then we know that the efficiency score of DMU_{j*} is the unity due to **Lemma 4.1**.

(ii) If for $j=1,\ldots,n$, we have

$$
\sum_{i=1}^{m}\omega_i^{\;0}x_{ij}-\sum_{r=1}^{s}\mu_r^{\;0}y_{rj}>0.
$$

Let $\alpha=\min_{1\le j\le n}\left(\dfrac{\sum_{i=1}^{m}\omega_i^{\;0}x_{ij}}{\sum_{r=1}^{s}\mu_r^{\;0}y_{rj}}\right)=\dfrac{\sum_{i=1}^{m}\omega_i^{\;0}x_{ij_*}}{\sum_{r=1}^{s}\mu_r^{\;0}y_{rj_*}}$, then we have $\alpha>1$ and $\sum_{i=1}^{m}\omega_i^{\;0}x_{ij}-\sum_{r=1}^{s}\alpha\mu_r^{\;0}y_{rj}\ge 0, j=1,\ldots,n$, where $\omega_i^{\;0}\ge\varepsilon>0,\alpha\mu_r^{\;0}\ge\varepsilon>0$. So, $\omega_i^{\;0},\alpha\mu_r^{\;0}$ are the feasible solutions of **Model 4.9**. But the inequality $\sum_{r=1}^{s}\alpha\mu_r^{\;0}y_{rj}>\sum_{r=1}^{s}\mu_r^{\;0}y_{rj}$ conflicts with the fact that $\omega_i^{\;0},\mu_r^{\;0}$ are optimal solutions. So, (ii) is impossible.

Here we know there are at least one DMU such that its efficiency score is the unity in the **Model 4.9**. Thus we have proved our conclusion.

Below we further show that PS may not be convex or free-disposable. Let $(a,b)\in R_+^2$ and consider all the vectors bigger than it:

$$S(a,b)=\{X=(x_1,x_2)\in R_+^2:(x_1,x_2)\succ(a,b)\}$$

It follows from the definition 4.2

$$S(a,b)=\{X=(x_1,x_2)\in R_+^2:\frac{a}{x_1}+\frac{b}{x_2}\le 2\}$$

This is an area surrounded by the hyperbolic line: $(ax_2+bx_1=2x_1x_2)$ (with asymptotic lines $(x_1=\frac{a}{2},x_2=\frac{b}{2})$. Now let the inputs be (1,2) and (2,1), consider their convex combinations $\{(2-\lambda,1+\lambda):0\le\lambda\le 1\}$. We can construct

$AS=\{X=(x_1,x_2)\in R_+^2:\frac{2-\lambda}{x_1}+\frac{1+\lambda}{x_2}\le 2,0\le\lambda\le 1\}$ as follows:（**non**-shadowed area with doted boundary）:

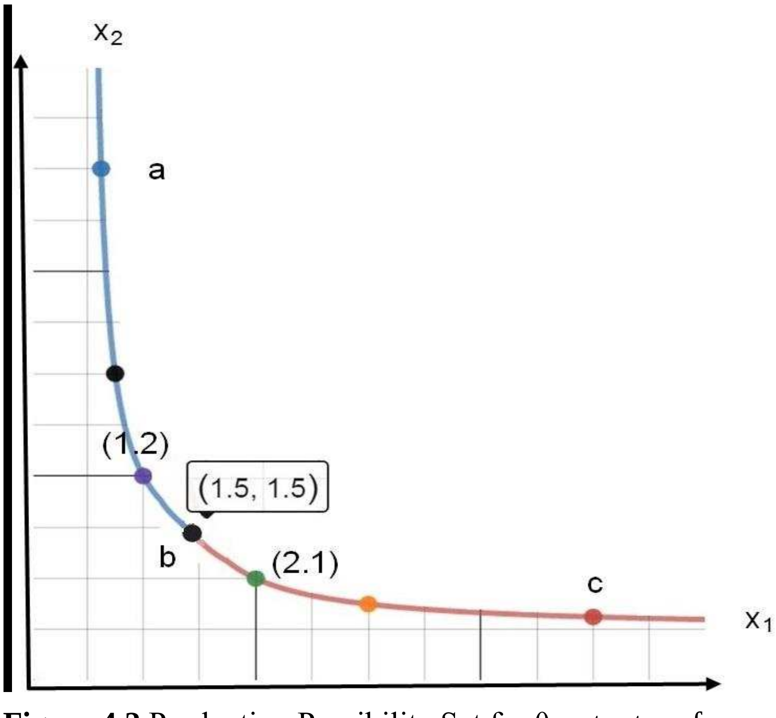

Figure 4.2 Production Possibility Set for θ-output preference

Let us note that all the hyperbolic lines from the convex combination pass the point: (1.5,1.5), thus the set AS is non-convex (only) around that point. Furthermore AS is also non free-disposable by noting this fact: let（a, b）be a point on the boundary of AS, then usually the set S(a, b) will contain points outside AS，as usually the boundary of S(a, b) differs from that of AS. As there is no contradiction loop, it is free-disposable in the weak sense defined in Section 3.1. Now let us include outputs. Firstly, in VRS case, let the output be the unity. Then it is clear that PS is a cylinder based on AS, between the planes y=0 and y=1. Thus it is non-convex and is non free-disposable. Thus we have given examples showing that the PS set of Model 2.3 may not be convex or free disposable. However, it is clear that PS set of Model 2.3 is better than that of Model 2.4 in the sense that it is only locally non-convex and it is weakly free-disposable. The above discussions hold for the case of CRS where PS is a cone starting from zero and the section cut by y=t is an area whose boundary is formed by the hyperbolic lines $t[(2-\lambda)\,y_2+(1+\lambda)\,y_1]=2y_1y_2$, which all pass $t^{1/3}(1.5,1.5)$.

In what follows we will examine the PS sets of Models 2.5-2.6.

4.3 Total slack preferences

Firstly, let us define the following preference: Total output slack preference:

$$(X_1, Y_1) \succ (X_2, Y_2) \Leftrightarrow \exists\ S^+, S^-, s.t.\ X_1 + S^- = X_2,\ Y_1 - S^+ = Y_2$$
$$S^- \geq 0,\ S^+ free,\ \sum_{r=1}^{s} s_r^+ \geq 0.$$

Obviously this preference is transitive and linearly compatible. The PPS with the total output slack preference reads:

$$\mathbf{PS} = \{(\mathbf{X},\mathbf{Y}) | \sum_{\mathbf{j}=1}^{\mathbf{n}} \lambda_{\mathbf{j}} \mathbf{x}_{\mathbf{ij}} \leq \mathbf{x}_{\mathbf{i}};\quad \mathbf{i} = 1,\ldots,\mathbf{m},$$
$$\sum_{\mathbf{j}=1}^{\mathbf{n}} \lambda_{\mathbf{j}} \mathbf{y}_{\mathbf{rj}} - \mathbf{s}_{\mathbf{r}}^+ = \mathbf{y}_{\mathbf{r}};\quad \mathbf{r} = 1,\ldots,\mathbf{s},$$
$$\lambda \in \mathbf{S},\ \sum_{\mathbf{r}=1}^{\mathbf{s}} \mathbf{s}_{\mathbf{r}}^+ \geq 0;\quad \mathbf{s}_{\mathbf{r}}^+\ \mathbf{free}\}$$

which is convex and free-disposable. With assumed technology set PS to be CRS, we can write the DEA model 2.6 as:

$$\max 1 + \frac{1}{s}\left(\sum_{r=1}^{s} \frac{s_r^+}{y_{r0}}\right) : \{(X,Y) \in PS\}.$$

Similarly, we can define total input slack preference:

$$(X_1, Y_1) \phi (X_2, Y_2) \Leftrightarrow \exists \, \text{slacks } S^+, S^-, \text{s.t. } X_1 + S^- = X_2, Y_1 - S^+ = Y_2$$
$$S^+ \geq 0, S^- \text{ free}, \sum_{i=1}^{m} S_i^- \geq 0.$$

And the PPS with the total input slack preference reads:

$$\begin{aligned} \mathbf{PS} = \{(\mathbf{X},\mathbf{Y}) | \sum_{j=1}^{n} \lambda_j \mathbf{x}_{ij} + \mathbf{s}_i^- = \mathbf{x}_i; \quad \mathbf{i} = 1, \mathbf{L} \ , \mathbf{m}, \\ \sum_{j=1}^{n} \lambda_j \mathbf{y}_{rj} \geq \mathbf{y}_r; \quad \mathbf{r} = 1, \mathbf{L} \ , \mathbf{s}, \\ \lambda \in \mathbf{S}, \quad \sum_{i=1}^{m} \mathbf{s}_i^- \geq 0; \quad \mathbf{s}_r^- \ \mathbf{free}\} \end{aligned}$$

Then assuming technology set S to be CRS, we can write the DEA model 2.5 as:

$$\min 1 - \frac{1}{m} \left(\sum_{i=1}^{m} \frac{s_i^-}{x_{i0}} \right) : \{(X, Y) \in PS\}.$$

For Model 2.5 we have the following theorem as well:

Theorem 4.3: There will always be at least one DMU for the **Model 2.5** such that its efficiency score is the unity.

Proof: Let $\theta_i = \dfrac{x_{i0} - s_i^-}{x_{i0}}, \ i = 1, \Lambda \ , m$, we have $s_i^- = x_{i0} - \theta_i x_{i0}, \ i = 1, \Lambda \ , m$. Note that the conditions: $\theta_i \geq 0, \ i = 1, \Lambda \ , m$ are implied in the first set of constraints of Model 2.5. Thus, Model 2.5 is equivalent to

$$\min \quad g_1 = \frac{1}{m} \sum_{i=1}^{m} \theta_i$$
$$\begin{cases} \sum_{j=1}^{n} \lambda_j x_{ij} = \theta_i x_{i0}, \ \ i = 1, \Lambda \ , m \\ \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0}, \ \ r = 1, \Lambda \ , s \\ \sum_{i=1}^{m} (x_{i0} - \theta_i x_{i0}) \geq 0, \\ \lambda_j \geq 0, \quad j = 1, \Lambda \ , n \\ s_r^+ \geq 0, \ \ r = 1, \Lambda \ , s \\ \theta_i \geq 0, \ \ i = 1, \Lambda \ , m \end{cases} \tag{4.10}$$

where the constraint: $\frac{1}{m}\sum_{i=1}^{m}\theta_i \le 1$ is implicitly enforced in the objective function. It is easy to find that the feasible set of Model 4.10 is no larger than that of Model 2.3. Thus, the optimal value of the above model is no less than that of Model 2.3. Note that there will always be at least one efficient DMU for Model 2.3. Thus, according to Theorem 4.2 at least one DMU's efficiency score calculated by Model 2.5 is the unity.

However, we cannot prove the same conclusion for Model 2.6 although computational results from random samples indicate it seems to hold as well.

Therefore, from the view point of preference these two models are more reliable to use. It follows from the above discussions that Model 2.3-2.6 are different from the standard ones in that they use different preferences either for input or output comparisons in their PPS. These preferences compare the performance of DMUs not in individual components of inputs or outputs (like Pareto) but in a sense of total performance. We will compare those models in the next section.

5. Example and Case Study

In this section, we intend to carry out some case studies to evaluate the efficiencies of 18 Chinese coastal cities and 15 basic research institutes in Chinese Academy of Sciences (CAS) in 2006, to compare the DEA models with the Russell measures. The first example comes from Zhu (1996), with the data given below:

Table 5.1 Data of 18 Chinese coastal cities

DMU	Cities	Input1	Input2	Output1	Output2	Output3
1	Dalian	2874.8	16738	160.89	80800	5092
2	QingHuangDao	946.3	691	21.14	18172	6563
3	TianJin	6854	43024	375.25	144530	2437
4	Qingdao	2305.1	10815	176.68	70318	3145
5	YanTai	1010.3	2099	102.12	55419	1225
6	WeiHai	282.3	757	59.17	27422	246
7	ShangHai	17478.6	116900	1029.09	351390	14604
8	LianYunGang	661.8	2024	30.07	23550	1126

9	NingBo	1544.2	3218	160.58	59406	2230
10	WenZhou	428.4	574	53.69	47504	430
11	GuangZhou	6228.1	29842	258.09	151356	4649
12	ZhangJiang	697.7	3394	38.02	45336	1555
13	BeiHai	106.4	367	7.07	8236	121
14	ShenZHen	4593.3	45809	116.46	56135	956
15	ZhuHai	957.8	16947	29.2	17554	231
16	ShaTou	1209.2	15741	65.36	62341	618
17	XiaMen	972.4	23822	54.52	25203	513
18	HaiNan	2192	10943	25.24	40267	895

We apply Models 2.3-2.6 to the above data and the results are as follows:

Table 5.2 Comparison of efficiency scores in different DEA models

DMU	**Model 2.3**		**Model 2.4**		**Model 2.5**		**Model 2.6**	
	Score	**Rank**	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**
1	0.3045	10	0.3900	8	0.3045	10	0.3900	9
2	1	1	0.4775	3	1	1	0.4775	7
3	0.1954	14	0.1441	15	0.1954	14	0.1441	15
4	0.3628	8	0.4575	5	0.3628	8	0.4575	8
5	0.6296	5	0.4769	4	0.6296	5	0.5385	4
6	1	1	0.3631	10	1	1	0.7990	2
7	0.2382	12	0.3114	11	0.2382	12	0.3114	11
8	0.3475	9	0.3819	9	0.3475	9	0.3819	10
9	0.6491	4	0.5438	1	0.6491	4	0.5510	3
10	1	1	0.4132	7	1	1	1	1
11	0.2071	13	0.2825	12	0.2071	13	0.2825	12
12	0.4864	7	0.5234	2	0.4864	7	0.5234	6
13	0.5160	6	0.4489	6	0.5160	6	0.5371	5
14	0.0848	18	0.0839	18	0.0848	18	0.0839	18
15	0.1039	17	0.0982	17	0.1039	17	0.0982	17

16	0.2592	11	0.2090	14	0.2592	11	0.2508	13
17	0.1680	15	0.2100	13	0.1680	15	0.2103	14
18	0.1216	16	0.1171	16	0.1216	16	0.1171	16

The correlations among the results are as follows:

Table 5.3 Pearson's Correlation Coefficients of scores

Models	Model2.5&2.6	Model 2.3&2.5	Model 2.4&2.6	Model 2.3&2.4
Correlation	0.9010	1	0.7487	0.6939

It follows from the correlation table that all the Models except Model 2.4 have produced highly correlated results, which is consistent with our theoretical analysis.

The other case study comes from a real evaluation of Chinese Academy of Sciences. Since the Pilot Project of Knowledge Innovation (KIPP) in 1998 in CAS, institutes' evaluation has become increasingly important and the requirements for the evaluation process have been diversified. The main feature of the institutes' evaluation at the present is the establishment of the Comprehensive Quality Evaluation (CQE) system, which is a comprehensive quality evaluation method. It uses fundamentally policy-oriented evaluation as the pivot, supported by quantitative monitoring. The CAS began to introduce the CQE in 2005. Under the framework of CQE, inputs and outputs of basic research institutes in CAS are monitored using several quantitative indicators. In Liu, et al. (2011), DEA based evaluation method was tested to explore the possibility of replacing the existing weighted sum methods used in CQE with a DEA based method to avoid the controversy of weights selection in the assessment. The results were compared with those obtained in CQE. The conclusion is that with suitable DEA models, they are consistent.

Here, we use the same inputs and outputs data set as in Liu et al. (2011) for DEA evaluation of the basic research institutes in CAS. The data are as follows, which come from quantitative monitoring report in 2006 in CAS, see Table 5.4 for details.

Table 5.4 Inputs and outputs of basic research institutes of CAS in 2006

DMU	**Inputs**		**Outputs**			
	Staff	**Res. Expen.**	**SCI Pub.**	**High Pub.**	**Grad. Enroll.**	**Exter.**

						Fund.
Unit1	380	59,880	201	28	386	35,368
Unit2	418	79,910	480	196	354	69,763
Unit3	68	13,150	78	72	57	5747
Unit4	1105	92,710	153	45	642	49,074
Unit5	248	18,920	68	18	165	13,801
Unit6	828	134,240	167	64	229	73,748
Unit7	481	52,460	38	13	136	32,797
Unit8	493	40,840	94	6	115	12,743
Unit9	198	23,110	43	16	79	15,964
Unit10	243	32,580	42	11	48	20,731
Unit11	553	62,100	156	34	105	67,927
Unit12	347	49,510	64	8	190	31,616
Unit13	445	78,280	440	162	529	62,448
Unit14	260	27,530	113	23	137	33,952
Unit15	304	59,450	94	19	263	70,015

*Value of Res. Expen. and Exter. Fund. are in RMB thousand.

Next, we will provide details of computational results of the evaluation of efficiency of basic research institutes in CAS in 2006 using Model 2.3 and Model 2.4. The results are presented in the following table.

Table 5.5 The efficiency scores of basic research institutes using Model 2.3&2.4

DMU	Model 2.3				Model 2.4					
	θ_1	θ_2	Score	Rank	θ_1	θ_2	θ_3	θ_4	Score	Rank
Unit1	0.8545	0.9539	0.9042	7	1.7671	11.7093	0.6724	0.7399	0.2687	7
Unit2	1	1	1	1	0.9875	2.2323	0.9785	0.5006	0.8513	2
Unit3	1	1	1	1	1.0011	0.4428	1.0103	1.9748	0.9031	1
Unit4	0.4887	1.0247	0.7567	9	3.5942	11.280	0.6260	0.8256	0.245	9
Unit5	0.5597	1.2905	0.9251	5	1.6504	5.7551	0.4970	0.5991	0.4705	4
Unit6	0.4201	0.5048	0.4625	13	4.7680	11.484	2.5409	0.7955	0.2042	10
Unit7	0.3118	0.5518	0.4318	14	8.1887	22.094	1.6720	0.6991	0.1225	13
Unit8	0.1962	0.4167	0.3065	15	2.5771	37.268	1.5393	1.4007	0.0935	15
Unit9	0.4117	0.6670	0.5393	11	3.1879	7.9084	1.2680	0.6327	0.3078	6
Unit10	0.3878	0.5636	0.4757	12	4.6012	16.216	2.9421	0.6868	0.1636	12
Unit11	0.5682	0.9845	0.7764	8	2.3612	10.000	2.5636	0.3995	0.261	8
Unit12	0.5168	0.6656	0.5912	10	4.5886	33.885	1.1295	0.6844	0.0993	14
Unit13	1	1	1	1	1.0553	2.6457	0.6414	0.5478	0.8179	3
Unit14	0.6442	1.1782	0.9112	6	1.4451	6.5537	0.8710	0.3544	0.4337	5
Unit15	1	1	1	1	3.7096	16.941	0.9689	0.3670	0.1819	11

It is clear that the scores of the two models are very different. The differences in rankings are smaller but still significant. These will be discussed further below. The table shows one of the advantages to use the above models is that the performance differences in each component of the inputs or outputs can be clearly seen.

We now compare Models 2.1-2.6 with a DEA model in Liu et al. (2011). When the standard DEA models are used in this data set, it was found that too many efficient DMUs were produced for an effective evaluation. Thus in the work of Liu et al. (2011), a DEA Model 15 was specially designed for the evaluation, where some of the data are allowed to have direct substitutions, and was recommended to be used in future CAS research evaluation.

Table 5.6 Comparisons of efficiency scores in different DEA models

DMU	Model 2.1		Model 2.2		Model 2.3		Model 2.4		Model 15		Model2.5		Model2.6	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
Unit1	0.9042	7	0.4181	8	0.9042	7	0.2687	7	0.561	7	0.9042	7	0.3031	9
Unit2	1	1	1	1	1	1	0.8513	2	1	1	1	1	1	1
Unit3	1	1	1	1	1	1	0.9031	1	1	1	1	1	0.9031	5
Unit4	0.765	9	0.3743	9	0.7567	9	0.245	9	0.2751	14	0.7631	9	0.2607	10
Unit5	1	1	1	1	0.9251	5	0.4705	4	0.4088	10	1	1	0.5752	6
Unit6	0.4625	13	0.219	11	0.4625	13	0.2042	10	0.4145	9	0.4625	13	0.2174	11
Unit7	0.4318	14	0.1429	13	0.4318	14	0.1225	13	0.3056	13	0.4318	14	0.1415	13
Unit8	0.3065	15	0.0935	15	0.3065	15	0.0935	15	0.1937	15	0.3065	15	0.0935	15
Unit9	0.5393	11	0.3644	10	0.5393	11	0.3078	6	0.4087	11	0.5393	11	0.3624	8
Unit10	0.4757	12	0.1898	12	0.4757	12	0.1636	12	0.3868	12	0.4757	12	0.1881	12
Unit11	0.7764	8	0.4625	7	0.7764	8	0.261	8	0.5627	6	0.7764	8	0.4592	7
Unit12	0.5912	10	0.1211	14	0.5912	10	0.0993	14	0.4318	8	0.5912	10	0.1200	14
Unit13	1	1	1	1	1	1	0.8179	3	0.925	4	1	1	0.9591	4
Unit14	1	1	1	1	0.9112	6	0.4337	5	0.6685	5	1	1	1	1
Unit15	1	1	1	1	1	1	0.1819	11	1	1	1	1	1	1

It is clear that Russell input and output models have produced very similar results. It is also clear that Models 2.3-2.6 have more discrimination power than the standard Russell DEA models as expected.

The Pearson's correlation coefficients of efficiency scores among Models 2.1-2.6 are shown in Table 5.7.

Table 5.7 Pearson's Correlation Coefficients

Models	Model 2.1&2.3	Model 2.2&2.4	Model 2.3&2.5	Model 2.4&2.6	Model 2.1&2.5	Model 2.2&2.6
Correlation	0.9943	0.7786	0.9579	0.7526	1	0.9464

It is clear that all models except Model 2.4 have produced highly correlated results. In the following, the Pearson's correlation coefficients of efficiency scores among Model 2.3-2.6 and Model 15 are shown in Table 5.8.

Table 5.8 Pearson's Correlation Coefficients

Models	Model 2.3&15	Model 2.4&15	Model 2.5&15	Model 2.6&15
Correlation	0.8155	0.7468	0.8092	0.8135

From Table 5.8, we find that the Pearson's Correlation Coefficients between Models 2.3 and Models 2.5-2.6 and Model 15 are much higher than those between Models 2.4 and Model 15 respectively. Overall we conclude that Model 2.3 and Model 2.5-2.6 all seem to be able to produce consistent results in these empirical tests. As discussed above, Model 2.4 is quite different from all others in terms of convexity and free-disposability, and should be avoided in real applications.

6. Conclusion and Discussion

In this paper we have studied and compared several DEA models in the Russell measures from the viewpoint of convexity and free-disposability of their PPS, resulted from the preferences used in the PPS. It is found that the standard Russell DEA models (Model 2.1-2.2) and the classic DEA models with the radial measure both uses the same preference (Pareto) in the PPS so that they are essentially the same models, although the classic Russell DEA models are able to measure the non-radial part of efficiency. However it is found in our studies that the θ-input, θ-output preference models (Model 2.3-2.4), and the total input slack, output slack preference models (Model 2.5-2.6) use the preferences very different from the Pareto either for input or output comparisons in their PPS. In the Pareto preference input/output A is better than input/output B if the components of A are all better than those of B, while in those preferences, this only needs a weaker condition that some total sums of the componets A is better than those of B. Thus the emphasis of these models is not to measure the non-radial part of efficiency, but some compensable sums of the relative measures of input/output components. Thus one uses them whenever one wishes to compare some total outcomes of input/output components, rather than their individual components. We further show that the preference used in the PPS of Model 2.4 has contradiction circles, which cause theoretical difficulties in interoperating the results. We have also carried out empirical studies to compare these models and found that except the θ-output preference model (Model 2.4), they all produce consistent results. Thus we do not recommend using the θ-output preference model (Model 2.4). It is clear that the total input slack, output slack preference models are more reliable in the sense that their PPS are both convex and strongly free disposable.

Acknowledgements

The authors would like to thank the support of the National Natural Science Foundation of China (Nos. 71771082, 71371067, 11501326) and Hunan Provincial Natural Science Foundation of China (No. 2017JJ1012).

References

[1] Aparicio J and Pastor J.T (2013), A well-defined efficiency measure for dealing with closest targets in DEA, Appl.Math.Compu., 219 (17): 9142-9154(2013).

[2] Bogetoft P and Hougaard J(1999), Efficiency Evaluations Based on Potential (Non-Proportional) Improvements, Journal of Productivity Analysis, 12(3):

233-247.

[3] Chen Y, and Sherman D.H (2004), The benefits of non-radial vs. radial super-efficiency DEA: an application to burden-sharing amongst NATO member nations, Socio-Economic Planning Sciences, 38: 307-320.

[4] Coelli Timothy J, Prasada Rao DS, Christopher J. O'Donnell, George E. Battese (2005), An Introduction to Efficiency and Productivity Analysis (2nd Edition), Springer Science Business Media, LLC, New York.

[5] Cook W.D, Zhu Joe (2007), Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, Springer Science+Business Media, LLC, New York.

[6] Cooper W.W, Seiford L.M, Zhu Joe (2004), Handbook on Data Envelopment Analysis, Kluwer Academic Publishers, Boston.

[7] Cooper W.W, Seiford L.M, Tone K (2006), Introduction to Data Envelopment Analysis and Its Uses, Springer Science+Business Media, LLC, New York.

[8] Cooper W.W, Seiford L.M, Tone K (2007), Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software (2nd Edition), Springer Science+Business Media, LLC, New York.

[9] Cooper W.W, Huang Z, Li S.X, Parker B.R, Pastor J.T (2007), Efficiency aggregation with enhanced Russell measures in data envelopment analysis, Socio-Economic Planning Sciences, 41:1-21.

[10] Fare R, and Lovell C. A. K (1978), Measuring the Technical Efficiency of Production, Journal of Economic Theory, 19: 150–162.

[11] Levkoff S.E，Russell RR, Schworm W, (2012), Boundary problems with the "Russell" graph measure of technical efficiency: a refinement, Journal of Productivity Analysis, 37: 239-248.

[12] Liu W.B, Sharp J, Wu Z.M (2006), Preference, production and performance

in data envelopment analysis, Annals of Operations Research, 145: 105-127.

[13] Liu W.B, Zhang D.Q, Meng W, Li X.X, Xu F (2011), A study of DEA models without explicit inputs, Omega, 39: 472-480.

[14] Liu W.B, Meng W, Li X.X, Zhang D.Q (2010), DEA models with undesirable inputs and outputs, Ann Oper Res, 173: 177-194.

[15] Liu W.B, Zhou Z.B, Liu D.B, Xiao H (2015a), Estimation of portfolio efficiency via DEA, Omega, 52: 107-118.

[16] Liu W.B, Zhou Z.B, Ma C.Q, Liu D.B, Shen W.F (2015b), Two stage DEA models with undesirable input-intermediate-outputs, Omega, 56: 74-87.

[17] Liu P.L (1985), Multiple Decision Making, Plenum, New York.

[18] Mirsalehy A, Bakar MRA , Jahanshahloo G.R, Lotfi F.H, Lee L.S(2016), Centralized resource allocation for connecting radial and non-radial models, Journal of Applied Mathematics, 6: 1-12.

[19] Pastor J.T, Ruiz J.L, Sirvent I (1999), An enhanced DEA Russell graph efficiency measure, European Journal of Operational Research, 115: 596-607.

[20] Ray S.C (2004), Data Envelopment Analysis Theory and Techniques for Economics and Operations Research, Cambridge university press, New York.

[21] Russell RR, Schworm W (2009), Axiomatic foundations of efficiency measurement on data-generated technologies, Journal of Productivity Analysis, 31: 77-86.

[22] Schaefer H.H, and Wolff M.P (1999), Topological vector spaces, 2nd ed, New York: Springer..

[23] Seiford L.M and Zhu Joe (1998), Identifying excesses and deficits in Chinese industrial productivity (1953-1990): a weighted data envelopment analysis approach. Omega, 26: 279-296.

[24] Seiford L.M and Zhu Joe (2003), Context-dependent data envelopment analysis--Measuring attractiveness and progress, Omega, 31: 397-408.

[25] Tone K (2001), A slack-based measure of efficiency in data envelopment analysis, European Journal of Operational Research, 130: 498-509.

[26] Zhang D.Q, Li X.X, Meng W, Liu W.B (2009), Measuring the performance of nations at the Olympic Games using DEA models with different preferences, Journal of the Operational Research Society, 60: 983–990.

[27] Zhou P, and Fan L.W (2010), An extension to data envelopment analysis with preference structure for estimating overall inefficiency, Applied Mathematics and Computation, 216: 812-818.

[28] Zhou Z, Sun L, Yang W, Liu W, Ma C (2013), A bargaining game model for efficiency decomposition in the centralized model of two-stage systems, Computers & Industrial Engineering, 64: 103-108.

[29] Zhu Joe (1996), Data Envelopment Analysis with Preference Structure, The Journal of the Operational Research Society, 47: 136-150.

[30] Zhu Joe (1998), Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities, European Journal of Operational Research, 111: 50-61.

[31] Zhu Joe and Cook W.D (2007), Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, Springer Science+Business Media, LLC, New York.

[32] Zhu Joe (2009), Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver (2nd Edition), Springer Science+Business Media, LLC, New York.