行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※

※ 資料包絡法在所有資料可同時變動之下 ※

※ 其效率的穩健性分析 (I) ※

※ Robustness of Efficiency in Data Envelopment Analysis ※

※ for Simultaneous Changes in all Data ※

※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：■個別型計畫　□整合型計畫

計畫編號：NSC 90－2218－E－040－003－

執行期間： 90 年 8 月 1 日至 91 年 7 月 31 日

計畫主持人：賴慶祥　中山醫學大學通識教育處

計畫參與人員：孫永昌　台中師範學院教育測驗統計所(碩士生)

本成果報告包括以下應繳交之附件：

□赴國外出差或研習心得報告一份

□赴大陸地區出差或研習心得報告一份

□出席國際學術會議心得報告及發表之論文各一份

□國際合作研究計畫國外研究報告書一份

執行單位：中山醫學大學通識教育處

中 華 民 國 91 年 10 月 1 日

行政院國家科學委員會專題研究計畫成果報告

資料包絡法在所有資料可同時變動之下其效率的穩健性分析(I)

Robustness of Efficiency in Data Envelopment Analysis for Simultaneous Changes in all Data

(第一年計畫)

計畫編號：NSC 90-2218-E-040-003
執行期限：90 年 8 月 1 日至 91 年 7 月 31 日
主持人：賴慶祥　中山醫學大學通識教育處
計畫參與人員：孫永昌 台中師範學院教育測驗統計所(碩士生)

中文摘要

資料包絡法主要是從所有的決策單元中區分出有效的決策單元，這些有效的決策單元組合成效率平面，本計劃對於資料包絡法的基本模式提出新的敏感度分析研究。藉著增加投入或減少產出，我們考慮一有效決策單元在部份決策單元子集合的資料可變動的情況下的穩健性問題；藉由修正的資料包絡法基本模式，使被評估的決策單元不包含在參考集合中，以找出可維持效率的穩定區間。

至目前為止，討論敏感度分析的文獻僅限於變動一個決策單元的投入或產出，來研究效率變動的問題，或是可同時變動全部決策單元的投入或產出，但僅找出可變動範圍的上限。在本研究中，資料允許同時變動全部或部份決策單元的投入或產出，亦可變動一個或多個的投入或產出。本敏感度分析提出一非線性規劃的模式，其最佳解乃是投入/產出的穩定區間，並具有增加投入或減少產出的充分必要條件。

本研究使資料包絡法的敏感度分析問題更趨一般化，並提供一方法使非線性規劃模式的最佳解可藉由線性規劃模式求解來得到其近似值。

關鍵詞：資料包絡法、決策單元、效率平面、敏感度分析、穩健性、非線性規劃。

Abstract

Data envelopment analysis (DEA) separates efficiencies from non-efficiencies. The efficient decision making units (DMUs) define the efficient frontier in data envelopment analysis. The project develops a new sensitivity analysis approach for all basic DEA models, such as, CCR, BCC and additive models, when the data of all DMUs are varied simultaneously. We consider the robustness of an extreme efficient DMU by giving increase in inputs or giving decrease in outputs of a subset of DMUs, if the DMU remains efficient after the change. By means of modified DEA models, in which the extreme efficient DMU under evaluation is not included in the reference set, and the stability region for maintaining efficiency is founded.

The existing literature on sensitivity analysis deals with only change the inputs and/or outputs of one DMU while the others remain unchanged, or only the upper bound of variation is found for simultaneous changes in all the data. In our framework, data are allowed to vary simultaneously for any subset of DMUs and across different subsets of inputs and outputs. This sensitivity analysis is modeled as a non-linear programming whose optimal values yield a stability region of an extreme efficient DMU. Sufficient and necessary conditions are provided for upward variations of input values of activity and downward variations of output values of activity for some DMUs such that an extreme efficient DMU remains efficient.

This approach generalizes the usual

sensitivity analysis of DEA and provided an algorithm that using linear programming model to approximate the optimal solution of the non-linear model.

Keywords: Data Envelopment Analysis (DEA), Decision-Making Unit (DMU), Efficient Frontier, Sensitivity Analysis, Robustness, Non-linear programming.

1. Background and Objective

Since the original publication [1], DEA has become a popular method for analyzing the efficiency of various organization units [2]. Relying on a technique based on mathematical programming without introduction of any subjective or economic parameters (weights, prices, etc.), DEA separates efficient from inefficient decision making units (DMUs) and indicate the 'efficient peers' for each inefficient DMU.

One of the important topics in DEA is the sensitivity of DMUs. Charnes et al. [3] first investigated the sensitivity of single output variation on the CCR model by updating the inverse of the optimal basis matrix. Charnes and Neralic [4] used the same technique to explore the sensitivity of the additive model in DEA for a simultaneous change in all inputs and/or all outputs of an efficient DMU. Andersen and Petersen [5] proposed the 'extended DEA measure' (EDM) model for ranking the efficient units. The EDM model (is also called super-efficiency model) was widely applied in the DEA sensitivity analysis [6-9]. It is based on modifying DEA models in which the test DMU is excluded from the reference set. Charnes et al. [6,7] provided a 1-norm and ∞-norm to compute stability regions for efficiency classifications under additive model. Zhu [8] used the super-efficiency model to determine necessary and sufficient conditions for preserving efficiency of the efficient DMUs under the CCR ratio model when data of the test efficient DMU were changed. Seiford and Zhu [9] generalized the method to yield the entire stability region of the test DMU.

The above sensitivity analysis literatures deal with the situation that the data variations are only applied to the test DMU. However, possible data errors may occur for each DMU simultaneously or independently. Thompson et al. [10] utilized Strong Complementary Slackness Condition (SCSC) multipliers to analyze the stability of CCR efficiency when the data for all efficient and inefficient DMUs were simultaneous changed in opposite directions. Seiford and Zhu [11] used the super-efficiency models to consider the data changes in all DMUs simultaneously. Their discussion is based on a worst-case scenario in which the efficiency of the test DMU is deteriorating while the efficiencies of all other DMUs are improving. For each efficient DMU, a range of stability for preserving efficiency is calculated.

In reality, uncertain conditions may affect a subset of DMUs only rather than all of DMUs, i.e., the possible data errors occur on the affected DMUs only. In this research, we used the modified DEA models to study the stability of efficient DMUs when the data of a subset (including the test efficient DMU) of DMUs are changed simultaneously in the same direction. By means of extended versions of super-efficiency additive model, we propose a non-linear programming problem whose optimal values yield particular stability regions for the test DMU. Sufficient and necessary conditions for preserving the test DMU remains efficient with respect to the data changed type are provided. These results remain valid under other DEA models when data of a subset of DMUs change simultaneously.

The following section proposed the sensitive analysis for DEA additive models. Non-linear models are proposed for finding the stability region of a test efficient DMU. Sufficient and necessary conditions are proved in this research. A numerical example is presented in Section 3. Discussions and conclusions are presented in Section 4.

2. Sensitivity Analysis

2.1 Type of Data Change

In the research, we are interested in the stability of a specific extreme efficient DMU_O while the data of a particular subset of DMUs, including DMU_O, is changed. Since either an increase of any output or decrease of any input cannot worsen an efficient DMU, data of the subset of DMUs are changed by giving upward variations in inputs or giving downward variations in outputs. Let ***P*** and ***U*** denote the sets of indices of all the DMUs that its data are changed and unchanged, respectively.

$\boldsymbol{P} = \{ j \mid \text{data of } DMU_j \text{ are changed}\}$

$\boldsymbol{U} = \{ j \mid \text{data of } DMU_j \text{ are unchanged}\}$

Suppose that ***I*** and ***O*** denote the input and output subsets respectively in which we are interested. In our research, we consider the absolute variations of data according to the following expressions:

For $DMU_j, j \in \boldsymbol{P}$

$$\begin{cases} \hat{x}_{ij} = x_{ij} + \Delta, \Delta \geq 0, & i \in \mathbf{I} \\ \hat{x}_{ij} = x_{ij}, & i \notin \mathbf{I} \end{cases} \qquad (1)$$

and

$$\begin{cases} \hat{y}_{rj} = y_{rj} - \delta, \delta \geq 0, & r \in \mathbf{O} \\ \hat{y}_{rj} = y_{rj}, & r \notin \mathbf{O} \end{cases} \qquad (2)$$

For $DMU_j, j \in \boldsymbol{U}$

$\hat{x}_{ij} = x_{ij}$, for all i.

$\hat{y}_{rj} = y_{rj}$, for all r.

where ($\hat{}$) represents the adjusted data.

2.2 Increase of Inputs

Let us first consider the model for change data in inputs only.

$$\begin{aligned} \Delta^* = & \operatorname{Min} \Delta \qquad\qquad (M1) \\ \text{s.t.} & \sum_{j \in U} \lambda_j x_{kj} + \sum_{j \in P, j \neq o} \lambda_j (x_{kj} + \Delta) \leq x_{ko} + \Delta, \\ & \sum_{j \neq o} \lambda_j x_{ij} \leq x_{io}, \ i = 1,2,\cdots,m;\ i \neq k, \\ & \sum_{j \neq o} \lambda_j y_{rj} \geq y_{ro}, \ r = 1,2,\cdots,s, \\ & \sum_{j \neq o} \lambda_j = 1, \\ & \Delta, \lambda_j \geq 0, \ j \neq o. \end{aligned}$$

Suppose the optimization of (M1) is completed for a specific index k. Sufficient and necessary conditions for preserving efficiency of DMU_O are shown as following.

Theorem 1. Given an increase as (1) in the k^{th} input only, the extreme efficient DMU_O remains efficient if and only if $\Delta \in [0, \Delta^*]$, where Δ^* is the optimal value to (M1).

This theorem show that the optimal solution of (M1) provides the possible maximum increase value for each individual input to allow DMU_O to be efficient when the other inputs are held at constants. We turn to consider the case there are more than one of inputs changed simultaneously in the same value. Assume each input $i \in \boldsymbol{I}$ is increased by the same value Δ. We also want to determine the possible maximum increase value of those interested inputs. We consider the following model extended from (M1).

$$\begin{aligned} \Delta^* = & \operatorname{Min} \Delta \qquad\qquad (M2) \\ \text{s.t.} & \sum_{j \in U} \lambda_j x_{ij} + \sum_{j \in P, j \neq o} \lambda_j (x_{ij} + \Delta) \leq x_{io} + \Delta, i \in \mathbf{I}, \\ & \sum_{j \neq o} \lambda_j x_{ij} \leq x_{io}, \ i \notin \mathbf{I}, \\ & \sum_{j \neq o} \lambda_j y_{rj} \geq y_{ro}, \ r = 1,2,\cdots,s, \\ & \sum_{j \neq o} \lambda_j = 1, \\ & \Delta, \lambda_j \geq 0, \ j \neq o. \end{aligned}$$

If we assume the problem is feasible, then the theorem is easily derived.

Theorem 2. The extreme efficient DMU_O remains efficiency after the absolute change in inputs as (1) if and only if $\Delta \in [0, \Delta^*]$, where Δ^* is the optimal value to (M2).

2.3 Decrease of Outputs

Now, turning to consider the case of changing data in outputs only, we utilize the following model.

$$\delta^* = \text{Min } \delta \quad \text{(M3)}$$
$$\text{s.t.} \sum_{j\in U} \lambda_j y_{kj} + \sum_{j\in P, j\neq o} \lambda_j (y_{kj} - \delta) \geq y_{ko} - \delta,$$
$$\sum_{j\neq o} \lambda_j y_{rj} \leq y_{ro}, \; r = 1,2,\cdots,s; \; r \neq k,$$
$$\sum_{j\neq o} \lambda_j x_{ij} \geq x_{io}, \; i = 1,2,\cdots,m,$$
$$\sum_{j\neq o} \lambda_j = 1,$$
$$\delta, \lambda_j \geq 0, \; j \neq \text{o}.$$

Suppose (M3) is maximized for a specific index k. Sufficient and necessary conditions for preserving efficiency of DMU_o are shown as follows.

Theorem 3. Given a decrease as (2) in the k^{th} output only, the extremely efficient DMU_o remains efficient if and only if $\delta \in [0, \delta^*]$, where δ^* is the optimal value to (M3).

The theorem illustrates that optimal solution of (M3) provides the possible maximum decrease value for each individual output to allow DMU_o to be efficient when the other outputs are held at constants. We turn to consider the case there are more than one of outputs are changed simultaneously in the same value. Assume each output $r \in \mathbf{O}$ is decreased by the same value δ. We also want to determine the possible maximum decrease value of those interested outputs. We consider the model extended from (M3).

$$\delta^* = \text{Min } \delta \quad \text{(M4)}$$
$$\text{s.t.} \sum_{j\in U} \lambda_j y_{rj} + \sum_{j\in P, j\neq o} \lambda_j (y_{rj} - \delta) \geq y_{ro} - \delta, r \in \mathbf{O}$$
$$\sum_{j\neq o} \lambda_j y_{rj} \leq y_{ro}, \; r \notin \mathbf{O},$$
$$\sum_{j\neq o} \lambda_j x_{ij} \geq x_{io}, \; i = 1,2,\cdots,m,$$
$$\sum_{j\neq o} \lambda_j = 1,$$
$$\delta, \lambda_j \geq 0, \; j \neq \text{o}.$$

If we assume the problem is feasible, then the following theorem is easily derived.

Theorem 4. The extreme efficient DMU_o remains efficient after the absolute change in outputs as (2) if and only if $\delta \in [0, \delta^*]$, where δ^* is the optimal value to (M4).

3. Numerical Example

The data sėt contains 8 DMUs with two inputs and one output is listed in Table 1. Points of DMU_1, DMU_2, DMU_3, and DMU_4 are BCC efficient while points of DMU_5, DMU_6, DMU_7, and DMU_8 are inefficient. The input values of DMU_2, DMU_3, and DMU_6 are increased simultaneously while the others are unchanged.

Table 1 shows that the stability region of DMU_2 and DMU_3 are 1.33332 and 3.33333 if we change input x_1, 2 and 2.33333 if we change input x_2, and 0.79951 and 1 if change input x_1 and x_2 simultaneously. The stability region of DMU_3 is larger than the stability region of DMU_2. It implies that DMU_3 is more stable than DMU_2 while changing the value of inputs in DMU_2, DMU_3, and DMU_6 simultaneously.

Table 1. The stability region for changing inputs of 3 DMUs on a data set of 8 DMUs.

DMU	y_1	x_1	x_2	Stability regions^c
1	1	1	12	
2	1	2	6	Δ^*_1= 1.33332, Δ^*_2= 2.00000, Δ^*= 0.79951
3	1	4	3	Δ^*_1= 3.33333, Δ^*_2= 2.33333, Δ^*= 1.00000
4	1	12	1	
5	1	2	8	
6	1	7	4	
7	1	6	7	
8	1	5	4	

^{c}: Δ^*_1, Δ^*_2, and Δ^* are the stability regions corresponding to change input x_1, x_2, and change all inputs simultaneously. The error tolerance is given by $\varepsilon = 10^{-3}$.

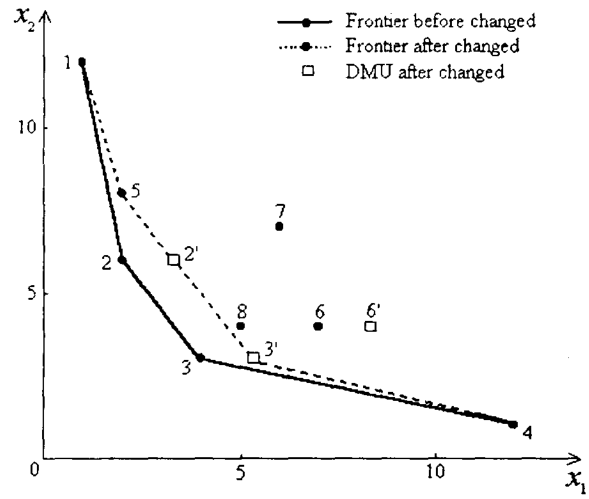

Figure1. The stability region of DMU_2 for increase of the value in input x_1 of DMU_2, DMU_3, and DMU_6 simultaneously.

Figure 1 presents the stability region of DMU_2 and frontiers before and after the change, while we changed the value of input x_1 in DMU_2, DMU_3, and DMU_6 simultaneously. The maximum increment of x_1 in DMU_2, DMU_3, and DMU_6 to preserve DMU_2 remaining efficient is approximated by $\Delta^*=1.33332$ (where the exact solution is 4/3). Therefore, DMU_2 fall in E' and can be expressed as the linear combination of DMU_3 and DMU_6 after the change.

4. Conclusion and Discussion

The paper presented a new DEA sensitivity approach referring to the non-linear models that may be considered as the extension of super-efficiency models [5]. The new sensitivity technique provides the stability of efficient DMUs by giving the data variations on a subset of DMUs. In contrast to the usual DEA sensitivity approaches whose data variations are considered either on the test DMU or on the allover DMUs, this approach proposed the generalized consideration that the uncertainty only affects a subset of DMUs. In reality, the data of DMUs may be varied in the same direction by the uncertainty. Sensitivity analysis enhances the fine quality of the final decision. Also, one can have the insight of the comparison between DMUs.

Depending on some results of DEA sensitivity analysis, we need to mention the specifics about set $\boldsymbol{P}$: (i) Since the efficient DMUs is always stable if an inefficient DMU is varied upward in inputs or downward in outputs, one may assume that set $\boldsymbol{P}$ contains the changed efficient DMUs only. For simplicity, $\boldsymbol{P}$ can be reset as $\boldsymbol{P}\cap\boldsymbol{E}$, (ii) The problem is similar to Zhu [8] if set $\boldsymbol{P}$ has only one element, say DMU_o. Zhu's method is generalized by this approach that $\boldsymbol{P}$ could contain DMU_o and other efficient DMUs, (iii) Suppose the reference set for solving DMU_o by the super-efficiency BCC models is denoted by Λ. The problem is also similar to Zhu [8] if $\boldsymbol{P}$ does not contain any one element of Λ. Thus, the main problem we employed is that the data changing on a set of DMUs, containing the test DMU and some of its reference points with respect to the super-efficiency DEA models. Zhu's approach [8] is the special case of ours.

Sufficient and necessary conditions are provided for upward variations of inputs and/or downward variations of outputs on a subset of DMUs simultaneously, such that an extreme efficient DMU remains efficient. Instead of effort on the difficulties for solving the non-linear problem, an algorithm applied the simpler linear model to approximate its optimal solution is also provided here. Our model obtained the same variations in all interested inputs and outputs, which is not necessarily true for the real-world applications. However, rescale all inputs and outputs suitably could be used to prevent this shortcoming. The above results also hold for other basic DEA models.

In this approach, we have assumed that the models we employed, that is (M1), (M2), (M3) and (M4), were always feasible. However, this assumption is not necessarily true. The conditions for infeasibility of super-efficiency models investigated by Seiford and Zhu [14] are useful for our proposed models. In accordance with Seiford and Zhu [11], (i) the infeasibility of modified super-efficiency models can be interpreted as stability of the efficiency classification of DMU_O with respect to the changes of corresponding inputs and/or outputs in a subset of DMUs, and (ii) the DEA measure of efficient DMUs for the BCC model is more stable than that in the CCR model. In fact, we could also employ the algorithm in Seiford and Zhu [9] to determine the whole stability region of the test DMU.

Although the stability region of a test efficient DMU for absolute changes in the data is identified, the values of data change are not necessarily the same within set ***P***. The possible future extensions of the research include: (i) changing data proportionally in a subset of DMUs, (ii) changing different scale in different input/output in a subset of DMUs, and (iii) other conditions for infeasibility, etc..

Appendix

Let's consider the following models.

$$\begin{aligned} \Delta(t) = \text{Min}\ &\Delta \qquad \text{(A1)} \\ \text{s.t.} &\sum_{j\in U}\lambda_j x_{ij} + \sum_{j\in P, j\neq o}\lambda_j (x_{ij}+t) \le x_{io} + \Delta, i\in \mathbf{I}, \\ &\sum_{j\neq o}\lambda_j x_{ij} \le x_{io},\ \ i\notin \mathbf{I}, \\ &\sum_{j\neq o}\lambda_j y_{rj} \ge y_{ro},\ r=1,2,\cdots,s, \\ &\sum_{j\neq o}\lambda_j = 1, \\ &\Delta, \lambda_j \ge 0,\ j\neq \text{o}. \end{aligned}$$

$$\begin{aligned} \Delta^{**} = \text{Min}\ &\Delta \qquad \text{(A2)} \\ \text{s.t.} &\sum_{j\in U}\lambda_j x_{ij} \le x_{io} + \Delta, i\in \mathbf{I}, \\ &\sum_{j\in U}\lambda_j x_{ij} \le x_{io},\ \ i\notin \mathbf{I}, \\ &\sum_{j\in U}\lambda_j y_{rj} \ge y_{ro},\ r=1,2,\cdots,s, \\ &\sum_{j\in U}\lambda_j = 1, \\ &\Delta, \lambda_j \ge 0,\ j\in U. \end{aligned}$$

We will apply the ablve two models and some properties related to these models to develop an algorithm for approximating the increase stability region Δ^*.

Property 1. Let $\Delta(t)$ be the optimal value to (16). Then, $\Delta(t)$ is non-decreasing in t.

Property 2. If $t< \Delta^*$. Then $t< \Delta(t) \le \Delta^*$.

Property 3. If $t> \Delta^*$. Then, $t\ge \Delta(t)$.

Property 4. $\Delta^{**} \ge \Delta^*$.

Property 5. If $\lambda_j(0)=0$ for all $j\in \boldsymbol{P}$. Then, $\Delta(0)= \Delta^*$.

An algorithm for approximating Δ^* is as follows.

Step 0. (Initialized) Solve (A2) to obtain Δ^{**}.

Step 0.1. If Δ^{**} is bounded, set upper bound $U=\Delta^{**}$. Otherwise, let $U=M$, where M is a given sufficient large number.

Step 0.2. Let lower bound $L=0$ and ε be the error tolerance for estimating Δ^*.

Step 1. Solve (A1) with $t=(U+L)/2$ to obtain $(\lambda_j(t), \Delta(t))$.

Step 1.1. If $\lambda_j(t)=0$ for all $j\in \boldsymbol{P}$ then set $\Delta^*= \Delta(t)$ and stop.

Step 1.2. If $t<\Delta(t)$ then set $L=\Delta(t)$. Otherwise, set $U=\Delta(t)$.

Step 2. If $(U-L)< 2\varepsilon$ then set $\Delta^*= (U+L)/2$ and stop. Otherwise, go to Step 1.

A bisection procedure is applied in the algorithm for convergence. If Δ^{**} is feasible in Step 0, Δ^* must be feasible and its approximation could be obtained. However, Δ^* may occur infeasible or its value exceeds a large number such that the test DMU tends to be stable while data is changed in a sufficient large scale. So, the upper bound U is set sufficient large value if Δ^{**} is infeasible in Step 0. In the real-world problems, one may identity that a test DMU is stable if the stability of inputs is infeasible or large enough relatively to the data range of all DMUs.

Reference

[1] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European Journal of Operational Research* 1978; 2: 429-444.

[2] Norman M, Stoker B. *Data envelopment analysis: the Assessment of Performance*, John Wiley, Chichester, England, 1991.

[3] Charnes A, Cooper WW, Lewin AY, Morey RC, Rousseau J. Sensitivity and Stability analysis in DEA. *Annuals of Operations Research* 1985; 2: 139-156.

[4] Charnes A, Neralic L. Sensitivity analysis of the additive model of data envelopment analysis. *European Journal of Operational Research* 1990; 48: 332-341.

[5] Andersen P, Petersen NC. A procedure for ranking efficient units in data envelopment analysis. *Management Science* 1993; 39: 1261-1264.

[6] Charnes A, Haag S, Jaska P, Semple J. Sensitivity of efficiency classifications in the additive model of data envelopment analysis. *International Journal of Systems Science* 1992; 23: 789-798.

[7] Charnes A, Rousseau J, Semple J. Sensitivity and stability of efficiency classifications in data envelopment analysis. *Journal of Productivity Analysis* 1996; 7: 5-18.

[8] Zhu J. Robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research* 1996; 90: 451-460.

[9] Seiford LM, Zhu J. Stability regions for maintaining efficiency in data envelopment analysis. *European Journal of Operational Research* 1998; 108: 127-139.

[10] Thompson RG, Dharmapala PS, Thrall RM. DEA sensitivity analysis of efficiency measures with applications to Kansas farming and Illinois coal mining. In: Charnes A, Cooper WW, Lewin AY, Seiford LM (eds.). *Data Envelopment analysis: Theory, Methodology and Applications*. Boston MA: Kluwer Academic Publishers, 1994, 393-422.

[11] Seiford LM, Zhu J. Sensitivity analysis of DEA models for simultaneous changes in all the data. *Journal of Operational Research Society* 1998; 49: 1060-1071.

[12] Charnes A, Cooper WW, Thrall RM. A structure for classify and characterizing efficiency and inefficiency in data envelopment analysis. *Journal of Productivity Analysis* 1991; 2: 197-237.

[13] Zhu J, Shen Z. A discussion of testing DMU's returns to scale. *European Journal of Operational Research* 1995; 81: 590-596.

[14] Seiford LM, Zhu J. Infeasibility of super-efficiency data envelopment analysis models. *INFOR* 1999; 37: 174-187.