MOST INFLUENTIAL OBSERVATIONS-SUPER EFFICIENCY

Shaik Khaleel Ahamed
Research Scholar, C.S.E.Dept
S.V.U.College of Engineering, S.V. University
Tirupati, A.P, 517501, India
khaleelska@gmail.com

Prof.MM Naidu
Professor, C.S.E. Dept
S.V.U.College of Engineering, S.V. University
Tirupati, A.P, 517501, India
mmnaidu@yahoo.com

Prof.C.Subba Rami Reddy
Professor, Statistics.Dept
S.V. University
Tirupati, A.P, 517501, India
csruma@yahoo.com

Abstract

To measure productive efficiency of competing firms a linear programming technique called Data Envelopment Analysis (DEA) is widely used. It requires first to adopt a production possibility set, that is axiomatic based and inefficient plans are projected on to the boundary of pp set, for the purpose of which radial and non-radial measures are used. These measures emanate from varying philosophies. The boundary of the pp set is sensitive to the extremely efficient firms, that are potentially influential observations. This study proposed a methodology to arrange the influential observations according to their influence on inefficient firms. The method goes beyond the super efficiency method introduced by Andersen and Petersen (1993), the purpose of which was to rank the efficient observations all of which posses unitary efficiency as revealed by Charnes, Cooper and Rhodes (1978) DEA formulation.

Keywords: Data Envelopment Analysis; Efficiency; Super Efficiency; Influential Observations.

I. INTRODUCTION

In a production environment where firms compete combining similar inputs to produce similar outputs, even if all the firms employ the same technology they differ by managerial efficiency, environment and size. Due to inefficiency contributed by different causes exogenous and endogenous, firms are found unsuccessful to employ potential inputs in order to produce potential outputs. That is firm is off its trajectory, so that it fails to realize maximum outputs produced with the aid of minimum inputs. If it replaces obsolete technology by modern technology, the switching over impact manifests in input substitution and/or output transformation. Between two firms the one that employs smaller inputs and produces grater outputs is more efficient, which is said to dominate the other one. The firms that are not dominated by other firms are the extremely efficient firms. The extremely efficient firms determine the boundary of all production possibilities. It is a practical frontier rather than hypothetical one. This boundary is determined such that, it is minimal boundary that envelops all the observations.

The extremely efficient observations are the influential observations, because deletion of any one of them contracts the set of production possibilities and the inefficient firms gain efficiency. This study provides a methodology to identify the most influential observations.

II. REVIEW OF LITERATURE

Data Envelopment Analysis is a linear programming technique that measures efficiency of firms which combine similar inputs to produce similar outputs. For this purpose one has to construct a production possibility set determined by observed input and output vectors. Construction of such a set requires certain axioms to be invoked.

a) Charnes, Cooper, and Rhodes (CCR, 1978) constructed a production possibility set that is consistent with the axioms of (i) Inclusion, (ii) Strong Disposability, (iii) Closure under ray expansion and contraction and (iv) Minimum extrapolation. The boundary of the production possibility set is choosen for efficiency measurement, which is determined by the extremely efficient firms. Two prominent approaches for efficiency measurement are input orientation and output orientation that are radial based. In the former approach holding outputs constant inputs are radially reduced to reach the boundary of the production possibility set. Consequently, potential inputs are estimated. In output orientation, holding input vector constant outputs are radially expanded till a point on the frontier is reached. The fundamental assumption of radial contraction or expansion is that the technology remains to be the same both at interior and boundary of the production possibility set, since neither input substitution nor output transformation is possible. The CCR formulation leads to the following linear programming problems:

(i) $$\max \sum_{r=1}^{s} v_r y_{ro}$$

such that $$\sum_{i=1}^{m} u_i x_{io} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$-\sum_{i=1}^{m} u_i x_{ij} + \sum_{r=1}^{s} v_r y_{rj} \le 0, \quad \text{j= 1, 2… n}$$

$$u_i \ge 0, v_r \ge 0$$

(ii) $$\min \sum_{i=1}^{m} u_i x_{io}$$

such that $$\sum_{r=1}^{s} v_r y_{ro} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$-\sum_{i=1}^{m} u_i x_{ij} + \sum_{r=1}^{s} v_r y_{rj} \ge 0, \quad \text{j=1, 2 …....n}$$

The linear programming problems (1) and (2) measure input and output efficiency respectively. These problems are constructed for n firms, each of which is represented by its own m componented input and s componented output vector.

x_{ij} : i^{th} input of the j^{th} firm

y_{rj}: r^{th} outptut of the j^{th} firm

x_{io}: i^{th} input of the firm whose efficiency is under evaluation.

y_{ro}: r^{th} output of the firm whose efficiency is under evaluation.

u_i and v_r are input and output weights respectively.

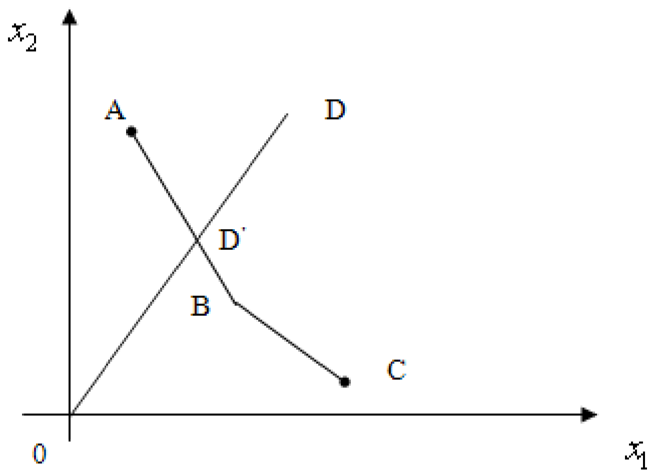


Figure (1)

Input efficiency

In the above figure we view a production process that employs two inputs to produce one output. The area lower bounded by the line segments AB and BC is the input set which is a cross section of general production possibility set. The line segment AB is generated by the linear combinations of the extreme points A and B. Similarly the line segment BC can be interpreted.

The producer D is inefficient and to attain input efficiency of one hundred percent his inputs are radially contracted to reach the boundary point D'. The radial input efficiency is

$$0 \le \frac{OD'}{OD} \le 1$$

The point D' is a linear combination of the extremely efficient points A and B. The decision making units A and B have similarities with D in terms of management and technology. The firms A and B influence D. The extreme points (A, B and C) are known as influential observations, since they influence inefficient decision making units. They lead and the inefficient firms follow.

With reference to problems (1) and (2), every point on the boundary of the technology set is viewed as linear combinations of input and output vectors.

$$x_i = \sum_{j=}^{n} \lambda_j x_{ij}, \ i \in M$$

$$y_r = \sum_{j=1}^{n} \lambda_j y_{rj}, \ r \in S$$

$$\lambda_j \ge 0, \ \ j \in N$$

b) Banker, Charnes and Cooper (1984) extended the CCR formulation, who imposed convexity on the production possibility set. The BCC production possibility set is a subset of CCR pp set. It admits variable returns to scale.

c) Andersen and Petersen (1993), for ranking purpose of most efficient decision making units introduced a method to identify the most influential observations, in the framework of CCR-DEA formulation. To assess the degree of influence of an efficient firm on inefficient firms, Andersen and Petersen (1993) estimated its super efficiency score which is larger than or equal to one under input orientation, but less than or equal to one under output orientation. Input super efficiency determines the capacity of efficient firm remain efficient under input expansion. Output super efficiency reveals the capacity of efficient firm to remain efficient under output contraction. To assess super efficiency of an extreme efficient firm, its input and output vectors are removed from the reference set, for the cause of which the production possibility set contracts. Consequently, it is assumed that larger is the super efficiency score, greater is the contraction of the production possibility set, attributed to the efficient firm under evaluation. The input super efficiency was used by Andersen and Petersen to rank the extremely efficient firms each of whose efficiency score is unity. Input super efficiency score of kth decision making unit can be assessed solving the following linear programming problem:

$$\lambda^{s} = \min \lambda$$

$$\text{such that } \sum_{\substack{j=1 \\ j\neq 1}}^{n} \lambda_j^k x_{ij} \leq \lambda x_{ik,}\ i \in M \qquad \ldots\ldots\ldots\ldots\ldots(3)$$

$$\sum_{\substack{j=1 \\ j\neq 1}}^{n} \lambda_j y_{rj} \geq y_{rk,}\ r \in S$$

$$\lambda_j \geq 0,\ j \in N$$

Since every feasible solution of (3) is a feasible solution of (2),

$$\lambda^{S} \geq \lambda^{T}$$

Super efficiency can be calculated only for extremely efficient firms for which $\lambda^{T} = 1$ always, thus $\lambda^{S} \geq 1$.

d) Wilson (1995) identified outliers following leave-one-out approach, and the search was in relation to efficient frontier, under exclusively input perspective or output perspective.
e) Simar (2003) suggests that a production plan shall be treated as an influential, if it is sufficiently influential under both the orientations (input and output). In order to determine a potential influential observation as an outlier one needs a threshold value with which the calculated metric is compared under constant returns to scale. The input and output efficiency scores are inversely related. Therefore inverse of threshold value of input orientation serves as threshold value of output orientation.
f) Stosic and Sampario de Souza (2003, 2005) proposed a method based on a combination of bootstrap and resampling schemes for automatic detection of influential observations, considered by them as outliers.
g) Tran et.al (2008) proposed a methodology for detecting influential observations. The method described below uses the number of positive intensity parameters and their sum as two metrics to identify influential observations.

   If $\lambda_j \geq 0$,then jth firm is an efficient peer of inefficient firm whose efficiency is under evaluation. We formulate the intensity parameter matrix, M_λ as follows:

$$M_\lambda = \begin{bmatrix} \lambda_1^{(1)} & \lambda_2^{(1)}\ldots\ldots & \lambda_j^{(1)}\ldots\ldots & \lambda_n^{(1)} \\ \lambda_1^{(2)} & \lambda_2^{(2)}\ldots\ldots & \lambda_j^{(2)} & \lambda_n^{(2)} \\ \lambda_1^{(k)} & \lambda_2^{(k)}\ldots\ldots & \lambda_j^{(k)}\ldots\ldots & \lambda_n^{(k)} \\ \lambda_1^{(n)} & \lambda_2^{(n)}\ldots\ldots & \lambda_j^{(n)}\ldots\ldots & \lambda_n^{(n)} \end{bmatrix}$$

$$\lambda^{T} = \min \lambda$$

$$\text{such that } \sum_{j=1}^{n} \lambda_j^{(k)} x_{ij} \leq \lambda x_{ik},\ i \in M$$

$$\sum_{j=1}^{n} \lambda_j^{(k)} y_{rj} \geq y_{rk},\ r \in S$$

$$\lambda_j \geq 0,\ j \in N$$

The above problem is solved for each k (=1, 2… n). kth row consists of the intensity parameters estimates obtained while kth firm is evaluated. The jth column consists parameter estimates of jth firm assigned by all other firms including itself. If the suffix j belongs to an inefficient firm, then every element in that column is zero. On the other hand, if the suffix j belongs to an efficient firm, then one or more of the elements in that column is/are positive. For each efficient firm we find intensity parameters (positive) count and intensity parameters sum.

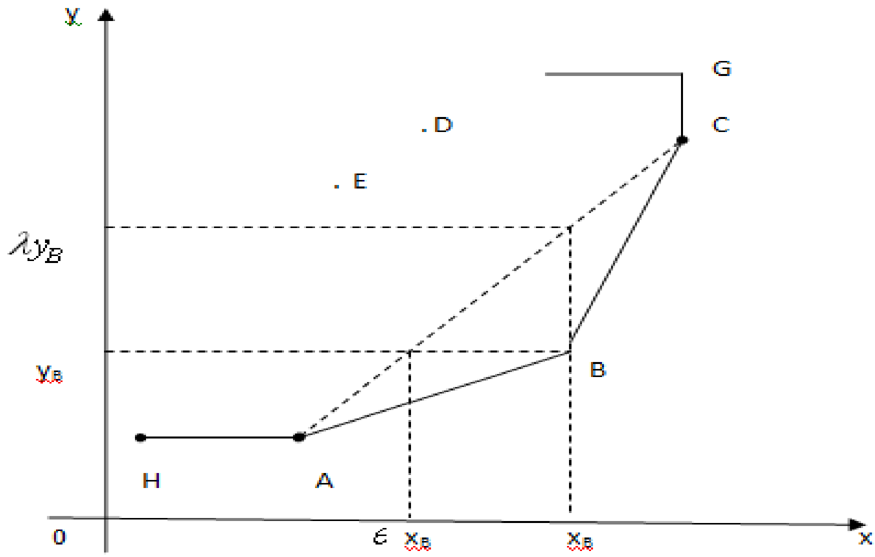
$$C_j = \sum_k I_k(j), \text{ where}$$

$$I_k(j) = \begin{cases} 1 if\, \lambda_j^{(k)} \geq 0 \\ 0 if\, \lambda_j^{(k)} = 0 \end{cases}$$

$$S_j = \sum_{k=1}^{n} \lambda_j^{(k)}$$

C_j and S_j can be used as metrics to determine how influential an extremely efficient firm is?

h) The leverage of an input-output observation to displace frontier is choosen as a metric to identify an outlier both in efficiency and inefficient perspectives. The leverage estimate is provided by super efficiency and super efficiency scores. If input orientation is pursued the threshold value of super efficiency scores is greater than unit, but less than unity for super inefficiency scores.



In the above inefficiency frontier is displayed which envelopes all the production plans from below. The inefficient frontier is determined by the firms H, A, B, C and G. B is extremely inefficient and its inefficiency score is unity. To assess its super efficiency, the production plan of B is removed from the reference set. The contracted production possibility set is bounded below by the inefficient frontier determined by H, A, C and G. the production plan of B is projected on to the displaced inefficiency frontier that indicated contraction of the production possibility set from below. The super efficiency score of firm B is θ that lies between zero and one.

$$0 \leq \theta_x \leq 1$$

If $(x_{B,}y_B)$ is flagged as an outlier in the input perspective, then the threshold value lies between θ and one. The super efficiency score under output orientation is larger than one.

$$\lambda_x \geq 1$$

and the consequent threshold value should also be larger than one. Chen and Johnson (2010) chose the inverse of input threshold value as output threshold value to identify the outlier

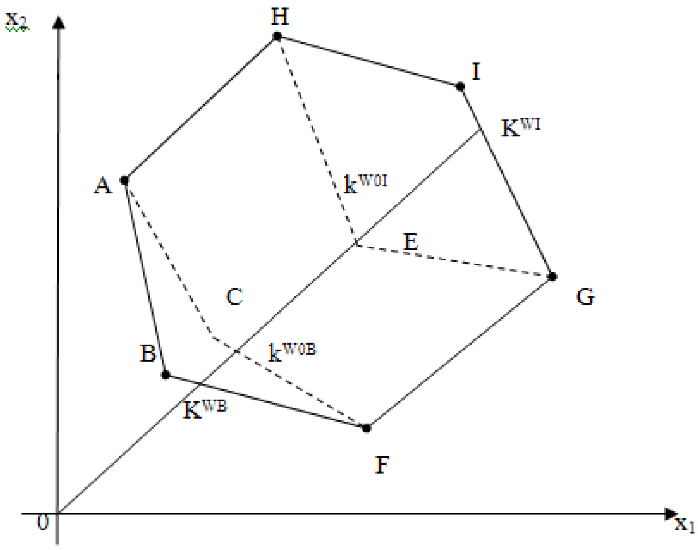
i) The efficient and inefficient frontiers bind the production possibility set from above and below respectively. The inefficient frontier violates the DEA axioms. As an alternative Chen and Johnson (2010) considered convex Hull, that satisfies the axioms of inclusion and convexity. The axiom of free disposability is withdrawn while this convex Hull is built.

$$T = \left\{ (x, y) : \sum_{j=1}^{n} \lambda_j x_j = x, \sum_{j=1}^{n} \lambda_j y_j = y, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j \in N \right\}$$

where $x \in R_m^+, y \in R_s^+$

The methodology developed to identify outliers is similar to the super efficiency evaluation proposed by Anderson and Petersen (1993). The leverage of an efficient DMU to contract the production possibility set while its input vector and output vector are removed from the reference technology determine if the efficient DMU under evaluation is outlier or not.

Removal of free disposability axioms, removes weak efficient subset of the DEA production possibility set from reference technology.



The above figure conveys that there are nine decision making units combining two inputs x_1 and x_2 to produce the same input level y.

$$\sum_{j=1}^{n} \lambda_j y_j = y \sum_{j=1}^{n} \lambda_j, \qquad \text{since } y = y_j, \forall j$$

$$= y$$

The constraint $\sum_{j=1}^{n} \lambda_j y_j = y$ is redundant. The convex hull is constructed by six decision making units ABFGIH. DMU_k is inefficient. The input vector of DMU_k can be radially contracted to k^{WB}, where 'WB' refers 'with B'. The input vector of DMUB is removed from the reference technology, consequently, the contraction measure k^{WB} is obtained, where 'WOB refer to 'without B'. The ability of B to contract the frontier is measured by Inner boundary shift = $\left[\frac{OK^{WOB}}{OK} - \frac{OK^{WB}}{OK} \right] \geq 0$

Similar measure can be obtained scaling up the DMU_K towards the frontier with and without DMU_I in the reference set.

$$\text{Outer boundary shift} = \left[\frac{OK^{WI}}{OK} - \frac{OK^{WOI}}{OK} \right] \geq 0$$

The outer boundary shift measures the power to contract convex hull from above.

$$\phi_0^N = \min \left\{ \phi : \sum_{j=1}^{n} \lambda_j x_j = \phi x_0, \sum_{j=1}^{n} \lambda_j y_j = y_0, \lambda_j \geq 0, j \in N \right\}$$

ϕ_0^N is radial measure and the movement is towards to efficient boundary, in the direction of input origin.

$$\prod_0^N = \max \left\{ \prod : \sum_{j=1}^{n} \lambda_j x_j = \prod x_0, \sum \lambda_j y_j = y_0, j \in N \right\}$$

$\prod_0^N$ is radial measure and the movement is towards inefficient boundary, in opposite direction from the input origin.

$$\phi_0^N \le 1, \prod_0^N \ge 1$$

The length of the ray, connecting efficient and inefficient boundary, involving (x_0,y_0).

$$\left\| \prod_0^N x_0 - \phi_0^N x_0 \right\| = \left(\prod_0^N - \phi_0^N \right) \left\| x_0 \right\|$$

To make this free from

units of measurement, $\dfrac{\left(\prod_0^N - \phi_0^N \right) \left\| x_0 \right\|}{\left\| x_0 \right\|} = \prod_0^N - \phi_0^N$

Let R be a point that determined the convex hull whose efficiency score or inefficient score is unity ($\prod_R^N = 1$ (or) $\phi_R^N = 1$)

To measure the ability of R to contract the distance evaluate by $\prod_0^N - \phi_0^N$, remove R from the reference set and obtain the contracted distance, $\prod_0^{N/R} - \phi_0^{N/R}$

$$\phi_0^N \le \phi_0^{N/R} \le 1$$

$$\prod_0^N \ge \prod_0^{N/R} \ge 1$$

The leverage of DMU_R on DMU_θ is measured by

$$(\prod_0^N - \phi_0^N) - (\prod_0^{N/R} - \phi_0^{N/R}) \ge 0$$

If R belongs to efficient boundary, then we have,

$$\prod_0^N = \prod_0^{N/R}$$

The leverage of R on DMU_0

$$\phi_0^{N/R} - \phi_0^N \ge 0$$

On the other hand, if R belongs to inefficient boundary, then

$$\phi_0^N = \phi_0^{N/R}$$

The leverage of R on DMU_0 is,

$$\prod_0^N - \prod_0^{N/R} \ge 0$$

$\prod_0^N = \phi_0^N = 1$ is possible, and in this case $\prod_0^{N/R} = \phi_0^{N/R} = 1$,which leads to

$(\prod_0^N - \phi_0^N) - (\prod_0^{N/R} - \phi_0^{N/R}) = 0$, implying that R has no leverage on DMU_0.

$$(\prod_0^N - \phi_0^N) - (\prod_0^{N/R} - \phi_0^{N/R}) = \underbrace{\left(\prod_0^N - \prod_0^{N/R} \right)}_{upperboundaryshift} + \underbrace{\left(\phi_0^{N/R} - \phi_0^N \right)}_{lowerboundaryshift}$$

This method can similarly be extended to the case of output orientation.

The impact measurements of R on all the other DMUs shall be summarized to measure its overall impact. Wilson (1995) used total value and the average number of individual influences, where the number of DMUs that are affected are also of interest.

III. NEW METHOD

If the input and output vectors of extremely efficient firms are deleted from the reference set, the relevant technology set contracts. As a result the efficiency scores of certain decision making units improve and for the rest of inefficient firms the efficiency scores remain to be the same. This method identifies such efficient firms whose deletion from the reference set leads to the significant contraction of the production possibility set as the most influential observation.

The super efficiency score is a single measurement concerned with the contracted space. However, we seek more measurements and they are available.

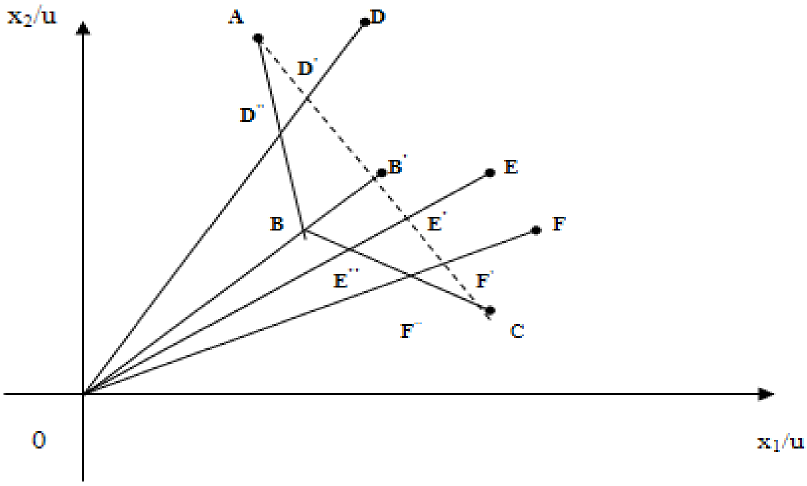


Figure (2)

Super efficiency

The input isoquant is determined by the extremely efficient decision making units A, B and C. The decision making units D, E, F are input inefficient. They consume more inputs than necessary. The input technical efficiencies of these firms are,

$$\lambda^T(D) = \frac{OD''}{OD}$$

$$\lambda^T(E) = \frac{OE''}{OD}$$

$$\lambda^T(F) = \frac{OF''}{OD}$$

If firm B is removed from the reference set, the fall out region of the production possibility set due to contraction is the dotted region. The new input frontier is the line segment AC and the modified efficiency scores of D, E and F are respectively, $\frac{OD'}{OD}$, $\frac{OE'}{OE}$ and $\frac{OF'}{OF}$. The super efficiency score of B is,

$$\frac{OB'}{OB} \geq 1$$

We have,

$$\frac{OD'}{OD} \geq \frac{OD''}{OD}$$

$$\frac{OE'}{OE} \geq \frac{OE''}{OE}$$

$$\frac{OF'}{OF} \geq \frac{OF''}{OF}$$

BB' that refers to super efficiency provides one measurement of the fallen out region of the technology set. The deviations which do not vanish provide additional measurements as below:

$$d_B = \frac{OD^{'}}{OD} - \frac{OD^{''}}{OD}$$

$$d_E = \frac{OE^{'}}{OE} - \frac{OE^{''}}{OE}$$

$$d_F = \frac{OF^{'}}{OF} - \frac{OF^{''}}{OF}$$

The deviation provided by super efficiency measurement is,

$$d_B = \frac{OB^{'}}{OB} - \frac{OB}{OB} = \frac{OB^{'}}{OB} - 1$$

The sum, $d=d_B+d_D+d_E+d_F$ provided a measure of the fallen out region of the production possibility set. The metric d, is viewed as area bound by a histogram with unit width class intervals, whose frequencies are the deviations.

To obtain these deviations two linear programming problems are solved. The first one is the CCR multiplier problem whose super efficiency version ignores the relevant constraint of the full problem. For k^{th} decision making unit we solve,

$$\lambda^k = \max \sum_{r=1}^{s} v_r y_{rk}$$

such that $\sum_{i=1}^{m} u_i x_{ik} = 1$ ……………………….. (P.1)

$$-\sum_{i=1}^{m} u_i x_{ij} + \sum_{r=1}^{s} v_r y_{rj} \leq 0, \text{ j=1, 2, 3……k-1, k+1..….n.}$$

$$u_i \geq 0, v_r \geq 0$$

$$\lambda^k \geq 1$$

Program (P.1) is solved only for the extremely efficient decision making units A, B, and C. The ratios,

$$\overline{d}_j = \frac{\sum_{r=1}^{s} \overline{v_r} y_{rj}}{\sum_{i=1}^{m} \overline{u_i} x_{ij}}$$

are calculated for all the inefficient decision making units, where $\overline{u_i}$ and $\overline{v_r}$ constitute the optimal solution of the program (P.1).

Let $\overline{x_{ij}} = \overline{d}_j x_{ij}$

for k^{th} firm the following CCR full multiplier problem is solved:

$$\overline{\lambda}^k = \max \sum_{r=1}^{s} v_r y_{rk}$$

$\sum_{i=1}^{m} u_i \overline{x_{ik}} = 1$ …………………..……..…. (P.2)

$$-\sum_{i=1}^{m} u_i \overline{x_{ij}} + \sum_{r=1}^{s} v_r y_{rk} \leq 0, \text{ j=1, 2…..…n}$$

$$u_i \geq 0, v_r \geq 0$$

Let $\overline{\overline{u}}_i$ and $\overline{\overline{v}}_r$ be the optimal solution of program (P.2)

The following ratios are calculated for inefficient firms:

$$\overline{\overline{d}}_j = \frac{\sum_{r=1}^{s} \overline{\overline{v}}_r y_{rk}}{\sum_{k=1}^{s} \overline{\overline{u}}_i x_{ik}}$$

$d_j = \overline{d}_j - \overline{d}_j \overline{\overline{d}}_j$ provides estimates of efficiency gains by inefficient decision making units due to the deletion of an efficient decision making unit from the reference technology.

$$d = \sum_{j \in I} (\overline{d}_j - \overline{d}_j \overline{\overline{d}}_j)$$

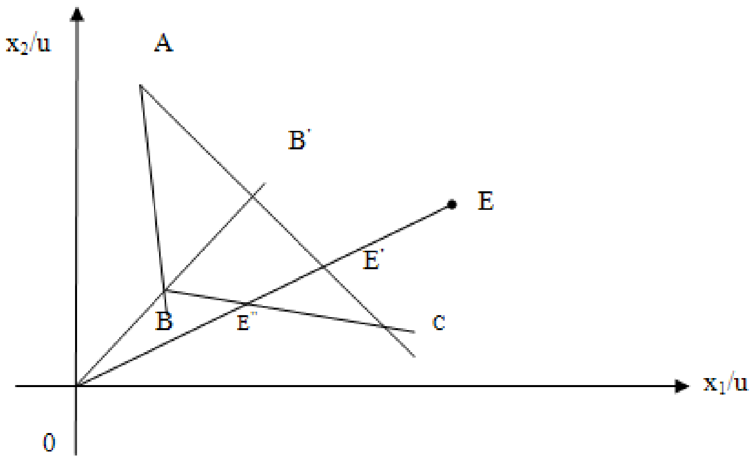
where $I = \left\{ j : \overline{d}_j - \overline{d}_j \overline{\overline{d}}_j \geq 0 \right\}$

IV. EMPIRICAL INVESTIGATION

19 public sector banks are considered to be firms or decision making units. Since these banks belong to the same sector the rules under which these function remains more or less to be same. Therefore, the environment that governs these banks is assumed to be uniform. The outputs are (i) interest income and (ii) non-interest income. The inputs are (i) deposits (ii) investments and (iii) number of employs. The data are collected from the reserve Bank of India Bulletins.

19 public sector banks are considered whose outputs and inputs are respectively, interest income and other income; Deposits, investments and number of employees. Assuming constant returns to scale to prevail, CCR (1978) problem is solved for each of the public sector banks. The optimization problem is CCR envelopment problem. Eight of these banks attained 100% efficiency score. These are denoted by DMU_2, DMU_4,DMU_8,DMU_{11},DMU_{12},DMU_{13},DMU_{14}, and DMU_{15}, where DMU refers to decision making unit, a commercial bank of the study.

For each of the efficient banks, its super efficiency problem is solved. This is done by deleting the input and output vector of CCR efficient DMU, whose efficiency is under evaluation. In the multiplier DEA this amounts to removal of the constraint of the efficient DMU for which the super efficiency problem is solved. The optimal solution produces n-1 efficiency ratios and one super efficiency ratio. The later exceeds unity, while each of the rest fall below or equal to one. For example, if input super efficiency is 1.5, then it implies that this bank remains to be efficient under input expansion upto a factor of 0.5.



In the above figure the line segments AB and BC constitute unit output, input isoquant. Bank B is extremely efficient. Deletion of B from the reference set results in a new frontier determined by the line segment AC. The super efficiency multiplier problem estimates the score equal to the ratio $\frac{OB'}{OB}$ which is larger than unity. For the inefficient bank E, efficiency ratio is, $\frac{OE^{'}}{OE}$. To project $E^{'}$ to $E^{''}$ we solve a multiplier problem (P.2) that yields the ratio, $\frac{OE^{''}}{OE^{'}}$

The product $\overline{(P.1)}$ of the two ratios yield,

$$\frac{OE^{'}}{OE} \times \frac{OE^{''}}{OE^{'}} = \frac{OE^{''}}{OE} \leq \frac{OE^{'}}{OE}$$

The deviation $\frac{OE^{'}}{OE} - \frac{OE^{''}}{OE^{'}}$ gives a measurement of the fallen out part of the production possibility set (the dotted region above). The super efficiency score, as usual, provides the one such deviation. The sum of all these deviations may be viewed as an area bound by the histogram of unit intervals whose frequencies being the magnitudes of the deviations. To examine if the deviational lengths are significantly different from zero, the arithmetic mean of all the deviations, ignoring the super efficiency deviation, since it is outlier, is tested against zero for statistical significance, using the Student's t-test. The test is right tailed. The results are interesting.

DMU2			**DMU4**			**DMU8**			**DMU11**		
$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation
0.9293	0.91861305	0.010687	0.803807	0.80099368	0.002813	0.753077	0.752996	0.010687	0.927432	0.926151	0.001281
1.0228	1.01369708	0.0228	0.869656	0.86748186	0.002174	0.731652	0.731574	7.8E-05	1	0.999391	0.000609
0.9049	0.90408559	0.000814	0.958154	0.95422557	0.003928	0.954987	0.954878	0.000814	0.874628	0.873054	0.001574
0.9989	0.99800099	0.000899	1.240614	1.23763653	0.2406	1	0.999886	0.000899	1	0.999441	0.000559
0.931	0.9190832	0.011917	0.908415	0.90596228	0.002453	0.765157	0.765149	0.011917	0.96095	0.960044	0.000906
0.9016	0.88978904	0.011811	0.834572	0.8310668	0.003505	0.843766	0.843674	0.011811	0.921577	0.92026	0.001317
0.8619	0.85207434	0.009826	0.702649	0.70173556	0.000913	0.521247	0.521194	0.009826	0.853219	0.852543	0.000676
0.9113	0.89845067	0.012849	0.913259	0.90768812	0.005571	1.228997	1.228868	0.229	0.951182	0.949365	0.001817
0.8406	0.82891566	0.011684	0.617462	0.61486866	0.002593	0.687734	0.687662	0.011684	0.827952	0.826465	0.001487
0.9878	0.9798976	0.007902	0.959654	0.95735083	0.002303	0.78285	0.782765	0.007902	0.984815	0.98373	0.001085
0.9989	0.98951034	0.00939	0.997441	0.99524663	0.002194	0.795277	0.795191	0.00939	1.011549	1.01059	0.011549
0.9866	0.97466214	0.011938	0.914293	0.90999582	0.004297	1	0.999889	0.011938	1	0.998348	0.001652
0.9989	0.98911078	0.009789	0.588761	0.58658258	0.002178	0.584186	0.584128	0.009789	0.92673	0.92502	0.00171
0.9523	0.93982487	0.012475	0.999953	0.99825308	0.0017	0.756544	0.756462	0.012475	1	0.999432	0.000568
0.999	0.9981009	0.000899	0.771273	0.771273	0	0.517139	0.517087	0.000899	0.913661	0.912685	0.000976
0.929	0.921568	0.007432	0.565757	0.56456891	0.001188	0.471905	0.471859	0.007432	0.858235	0.856982	0.001253
0.9653	0.95622618	0.009074	0.999942	0.9965422	0.0034	0.910967	0.910866	0.009074	0.984103	0.982867	0.001236
0.8357	0.83285862	0.002841	0.989093	0.9870159	0.002077	0.775907	0.775822	0.002841	0.903224	0.902732	0.000492
0.8802	0.87086988	0.00933	0.612784	0.61082309	0.001961	0.5656	0.565542	0.00933	0.842267	0.840976	0.001291
	am	0.008287		am	0.015045		am	0.008287		am	0.00168621
	sd	0.004182		sd	0.054635		sd	0.004182		sd	0.00242436
	root n	4.123106		root n	4.123106		root n	4.123106		root n	4.123106
	sd/root n	0.001014		sd/root n	0.013251		sd/root n	0.001014		sd/root n	0.00058799
	t-value	8.169426		t-value	1.135369		t-value	8.169426		t -value	2.86773558
	sum	0.174357		sum	0.28662		sum	0.367786		sum	0.032038

DMU12			**DMU13**			**DMU14**			**DMU15**		

$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation	$P.1$	$\overline{P.1}$	Deviation
0.953528	0.952595	0.000933	0.9213	0.919409	0.001891	0.7719	0.751522	0.020378	0.9293	0.867178	0.005322
1	0.999189	0.000811	1	0.997941	0.002059	0.8084	0.769435	0.038965	1.0228	0.9948	0.0052
0.799722	0.798941	0.000781	0.9249	0.923034	0.001866	0.7699	0.69676	0.07314	0.9049	0.840839	0.005161
0.871821	0.870964	0.000857	0.9753	0.973312	0.001988	1	0.8968	0.1032	0.9989	0.977998	0.006002
0.968461	0.967509	0.000952	0.8982	0.896343	0.001857	0.8968	0.879223	0.017577	0.931	0.887056	0.005444
0.955978	0.955042	0.000936	0.9008	0.898957	0.001843	0.816	0.8058	0.0102	0.9016	0.815893	0.005007
0.86661	0.86576	0.00085	0.8182	0.816493	0.001707	0.6925	0.670825	0.021675	0.8619	0.873141	0.005359
1	0.999189	0.000811	0.9469	0.944991	0.001909	0.8833	0.880915	0.002385	0.9113	0.764508	0.004692
0.909116	0.908231	0.000885	0.8573	0.85555	0.00175	0.6132	0.6132	0	0.8406	0.751488	0.004612
0.951179	0.950246	0.000933	0.9591	0.957126	0.001974	0.8754	0.825065	0.050336	0.9878	0.963685	0.005915
0.981173	0.980209	0.000964	0.9611	0.959115	0.001985	0.9328	0.887746	0.045054	0.9989	0.976308	0.005992
1.034155	1.033145	0.034155	1	0.997966	0.002034	0.8729	0.855529	0.017371	0.9866	0.876819	0.005381
1	0.999172	0.000828	1.0123	1.010228	0.0123	0.5557	0.537084	0.018616	0.9989	0.944304	0.005796
0.986398	0.985424	0.000974	0.8967	0.894831	0.001869	1.0003	0.979294	0.003	0.9523	0.93526	0.00574
0.856828	0.855992	0.000836	0.9542	0.952227	0.001973	0.6436	0.576215	0.067385	1.087724	0.999	0.087724
0.894442	0.893572	0.00087	0.9144	0.912513	0.001887	0.5261	0.499322	0.026778	0.929	0.929893	0.005707
0.9579	0.95696	0.00094	0.951	0.949053	0.001947	0.9222	0.880055	0.042145	0.9653	0.901368	0.005532
0.872927	0.872063	0.000864	0.7876	0.785964	0.001636	0.9875	0.967454	0.020046	0.8357	0.801779	0.004921
0.889479	0.888613	0.000866	0.8764	0.874598	0.001802	0.5867	0.56904	0.01766	0.8802	0.83905	0.00515
	am	0.002634		am	0.00243563		am	0.0313637		am	0.00971879
	sd	0.00763337		sd	0.00239127		sd	0.0269777		sd	0.01889428
	root n	4.123106		root n	4.123106		root n	4.123106		root n	4.123106
	sd/root n	1.02E-06		sd/root n	0.00057997		sd/root n	0.0065431		sd/root n	0.00458254
	t-value	867.30232		t-value	4.199604		t-value	4.7934427		t-value	2.12083174
	sum	0.050046		sum	0.046277		sum	0.595911		sum	0.184657

Note: Sum refers to sum of 19 deviations, t-test is based on 18 deviations, and Super Efficiency deviation is not included.

I. DMU_2 is extremely efficient. Its super efficiency score is 1.0228. This commercial bank remains to be radial input efficient under input expansion of 2.3 percent. The mean deviational length is statistically significant at 1% level of significance. The sum of all deviations that includes the super efficiency deviations is 0.1744 which implies that only 13.1 percent of the contracted region is accounted for by super efficiency.

II. DMU_4 is extremely efficient. Whose input super efficiency score is 1.2406 which implies that this commercial bank remains to be input technical efficient under input expansion of 24 percent. Its efficiency deviational sum is 0.28662, implying that 84 percent of production possibility set contraction is accounted for by super efficiency.

III. DMU_8 is extremely efficient. Its input super efficiency score is 1.229. This commercial bank remains to be efficient under input expansion of 2.3 percent. Its efficiency deviational sum amounts to 0.367786. 62 percent of the pp set contraction is accounted for by super efficiency.

IV. The super efficiency of DMU_{11} is 1.011549. This commercial bank is although marginally super efficient, the pp set contracted region unaccounted for by the super efficiency is about 64 percent.

V. DMU_{12} is extremely efficient. It is also marginally super efficient. It continues to be radial input technical efficient under input expansion upto 5 percent. Its efficiency deviational sum is 0.050046. 68 percent of production possibility set contraction is accounted for by super efficiency of the commercial bank.

VI. The super efficiency score of DMU_{13} is 1.0123. It is marginally super efficient. Its efficiency deviational sum is 0.046277 largest efficient banks. Only 27 percent of fallen out technology set is accounted for by the super efficient commercial bank.

VII. DMU_{14} is extremely efficient. Its super efficiency score is 1.003, implying that it is very marginally super efficient. However, its efficiency deviational sum is 0.595911. This commercial bank is least super efficient in its comparison to the other extremely efficient decision making units, its super efficiency accounts for only 5 percent of the contracted portion of the technology set. This commercial bank is the most influential observation that is revealed by the efficiency deviational sum which measures the contraction portion of the production possibility set.

VIII. The super efficient score of DMU_{15} is 1.087724, which implies that this commercial bank remains to be input technical efficient under input expansion of 8.8 percent. Its efficiency deviational sum is 0.1847. 20 percent of the fallen out production possibility set is accounted for by the super efficiency score of the commercial bank.

Table (2)

Commercial Bank	Efficiency Deviational sum	Super Efficiency Deviation	Contribution of Super Efficiency to the fallen out Production Possibility Set(percentage)
DMU_{14}	0.5959	0.003	0.5(0.005)
DMU_8	0.3679	0.229	62(0.6226)
DMU_4	0.2866	0.2406	84(0.8395)
DMU_{15}	0.1866	0.0877	48(0.4751)
DMU_2	0.1744	0.0228	13(0.1308)
DMU_{12}	0.05005	0.0342	68(0.6825)
DMU_{13}	0.0463	0.0123	27(0.2658)
DMU_{11}	0.032	0.0116	36(0.3605)

The Correlation Coefficient between super efficiency deviation and the deviational sum is found to be, R=0.2856

Its significance is tested at zero by Student's t-test.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = 0.73$$

The Correlation Coefficient is not significantly different from zero at 1 percent level of significance. Thus, a DMU with larger super efficiency ratio need not out right conclude that a DMU with the largest super efficiency score is the most influential observation.

V REFERENCES

[1] Andersen, P. and N. C. Petersen,(1993), "A Procedure for Ranking Efficient Units in Data Envelopment Analysis", Management Science, 39:1261-1264.

[2] Charnes, W.W. Cooper, Z.M. Huang and D.B. Sun, (1990), "Polyhedral cone-ratio DEA models with an illustrative application to large commercial bank", Journal of Economics, 46:73-91.

[3] Banker, Charnes and Cooper (1984), "Estimating Most Productive Scale Size Using Data Envelopment Analysis", European Journal Of Operations Research, 17:35-44.

[4] Charnes, A., Cooper W.W., and Rhodes, E., (1978), "Measuring the Efficiency of Decision-Making Units", European Journal of Operations Research, 2: 429-444.

[5] Chen and Johnson (2010), "A Unified model for detecting Outliers in DEA", Computers and Operations Research, Vol. 37:417-425.

[6] Daraio, C. and L. Simar (2003), " Introducing environmental variables in nonparametric frontier models: a probabilistic approach", Discussion paper 0313, Institute de Statistique, Universities Catholique de Louvain, Belgium.

[7] Johnson, A.L., Chen W.C., McGinnis, L.F.,(2008), "Internet-based benchmarking for warehouse operations". Working Paper, 2008.

[8] J.R. Doyle and R.H. Green (1994), "Efficiency and cross-efficiency in DEA: derivations, meanings and uses", Journal of Operational Research Society 45:567-578.

[9] J.H. Dula and B.L. Hickman (1997), "Effects of excluding the column being scored from the DEA envelopment LP technology matrix", Journal of Operational Research Society 48:1001-1012.

[10] J. Zhu (1996), "Robustness of the efficient DMUs in data envelopment analysis", European Journal of Operational Research 90:451-460.

[11] J. Zhu (2001), "Super-Efficiency and DEA sensitivity analysis", European Journal of Operational Research 129: 443-455.

[12] M. Halme and P. Korhonen (2000), "Restrciting weights in value efficiency analysis", European Journal of Operational Research 126:175-188.

[13] P.C. Pendharkar (2005), "A Data Envelopment Analysis-Based Approach for Data Preprocessing," IEEE Transactions on Knowledge & Data Engineering, Vol. 17, No. 10:1379-1388.

[14] R.G. Dyson and E (1988), "Thanassoulis, Reducing weight flexibility in data envelopment analysis", Journal of Operational Research Society 39:563-576.

[15] R. Green, J.R. Doyle and W.D. Cook (1996), "preference voting and project ranking using DEA and cross-evaluation", European Journal of the Operational Research 90:461-472.

[16] Stosic, B. and Sampaio de Sousa, M.C (2003), "Jackstrapping Dea Scores For Robust Efficiency Measurement", Series Texto para Discussão:291, Universidade de Brasília.

[17] Shaik Khaleel Ahamed, M. M. Naidu, C. Subba Rami Reddy (2015), "Outliers In Data Envelopment Analysis", International Journal of Computer Science and Security (IJCSS), Volume - 9 , Issue-3:164 – 173.

[18] S. Talluri and J. Sarkis (1997), "Extensions in efficiency measurement of alternate machine component grouping solutions via data envelopment analysis",IEEE Transactions on Engineering Management 44:27-31.

[19] Tran,N.M, Sheverly,G and Preckel,P (2010), " A New Method for detecting Outliers in DEA", Applied Economic Letters,17:313-316.

[20] T. R. Anderson, A. Uslu, and K. B. Hollingsworth (1998), "Revisiting extensions in efficiency measurement of alternate machine component grouping solutions via data envelopment analysis," Working paper.

[21] Timmer, C. Peter,(1971), "Using a probabilistic frontier production function to measure technical efficiency", Journal of Political Economy, 79:776-794.

[22] Wilson, P. W. (1995), "Protecting Influential Observations in Data Envelopment Analysis", Journal of Productivity Analysis, 4:27–45.