TWMS J. Pure Appl. Math., V.XX, N.XX, 20XX, pp.XX-XX

# RANKING DMUS BY BOOTSTRAPPING METHOD

SAID EBADI[1]

Abstract. The efficiency of Groups is determined by a comparison of their inputs and outputs. For multiple inputs and outputs such a comparison is not straightforward but can be accomplished using Data Envelopment Analysis. This is a technique based on linear programming which compares each unit with all the others and determines its efficiency in terms of other Groups with comparable inputs and outputs. This article suggests bootstrapping method for ranking measures of technical efficiency as calculated via Data Envelopment Analysis (DEA). This technique is employed to evaluate the efficiency of 9 sections of hospitals in Ardabil.
Keywords: Data Envelopment Analysis, Decision Making Unit, Efficiency, Bootstrapping.

AMS Subject Classification:90

## 1. Introduction

The Farrell [10] measure of technical efficiency for each decision making unit (DMU) within a sample of data can be calculated using two very different approaches for determining the production frontier. A nonparametric non-stochastic approach for measuring efficiency has been developed by Charnes, Cooper, and Rhodes [4] and Fare, Grosskopf, and Lovell [9], among others. Parametric stochastic models of efficiency measurement, beginning with the works of Aigner, Lovell, and Schmidt [1] , have also been introduced. Stochastic formulation of the original DEA models was introduced to incorporate possible uncertainty in inputs and/or outputs (e.g. Cooper, Deng, Huang, and Li [5]; Cooper, Seiford, and Tone [6]; Huang and Li [11]; Khodabakhshi and Asgharian [15]). Morita, Hiroshi, and Seiford [16] studied the robustness of efficiency results when input and output data are subject to stochastic measurement error, while Jess, Jongen, Neralic, and Stein [12] introduced a semi-infinite programming model in DEA to study a chemical engineering problem. More recently, stochastic input and output variations into DEA have been studied by; for example, Ebadi, Jahanshahloo, Monzeli and Aliev [7], Asgharian, Khodabakhshi, and Neralic [2], Khodabakhshi [14], and Khodabakhshi and Asgharian [15]. See, also, Kall [13] for discussions on linear programming programs.

Each of approaches has advantages and disadvantages relative to the other. Unfortunately, both the programming and econometric frontier methods share a common shortcoming. Using either approach, it is difficult to determine the statistical precision of the resulting efficiency scores. But the lack of measure of statistical precision for the efficiency scores limits their usefulness to decision makers.

This article shows how bootstrapping techniques can be used to construct measur for the efficiency scores produced by the linear programming approach to efficiency measurement. The bootstrap is a nonparametric approach to statistical inference. Alternatively, parametric or semi-parametric methods could be used to construct measur. The bootstrap was chosen because, like

[1]Department of Mathematics, Ardabil Branch, Islamic Azad University, Ardabil, Iran
e-mail: said_ebadisharafabad@yahoo.com
*Manuscript received XXXXXXXX 20XX.*

the linear programming approach itself, it is nonparametric and therefore does not impose any structure on the shape of the efficiency distributions. In addition to constructing measurs, the bootstrap can be used to calculate the bias of the original programming efficiency score. The article proceeds as follows. First, the programming approach to efficiency measurement is outlined. Our implementation of the bootstrap to establish statistical properties of the efficiency measure is then described. We then offer an illustration of this method by applying it to the 9 sections of hospitals in Ardabil.

## 2. DEA Background

DEA provides a measure of the efficiency of a DMU relative to other such units, producing the same outputs with the same inputs. The units to be compared may be enterprises, banks, schools, hospitals, etc. DEA is related to the concept of technical efficiency and can be considered as a generalization of efficiency measure. Furthermore, DEA possesses many advantages over techniques such as performance ratios and regression analysis, which makes it a vary suitable tool for management in a wide variety of industries.

Assume that there is a sample of $n$ DMUs, each producing an $s$-dimensional row vector of outputs $y$, from an $m$-dimensional row vector of inputs $x$. Technology governs the transformation of inputs into outputs; the reference technology relative to which efficiency is assessed is given by the input requirement set $L(y) = \{x : x \textit{ can produce } y\}$. Farrell's [10] input-based measure of technical efficiency for each observation $t = 1, ..., n$ is given by:

$$TE_t(x_t, y_t) = min\{\theta_t : \theta_t.x_t \in L(y_t)\}$$

that is, $t^{th}$ DMU's observed input vector $(x_t)$ is scaler $(0 \leq \theta_t \leq 1)$ until it is still just able to produce the observed level of output $(y_t)$. The solution, $TE_t = {\theta_t}^*$, gives the proportion of the $t^{th}$ DMU's actual input vector that is technologically necessary to produce its observed output vector given the best practice technology as revealed by the observed data. The vector ${x_t}^* = \theta_t^*.x_t$ would give the technically efficient (optimal) input vector for the $t^{th}$ DMU.

One way to calculate this measure of technical efficiency is by the following linear programming problem (Banker Charnes Cooper, BCC-Model) [3] once for each $DMU_t,\ t = 1, ..., n$ :

$$\begin{aligned} min \quad & \theta_t \\ st: & \\ & \lambda.Y \geq y_t \\ & \lambda.X \leq \theta.x_t \\ \sum_{t=1}^{n} & \lambda_t = 1 \\ & \lambda_t \geq 0 \quad t = 1, ..., n. \end{aligned} \tag{1}$$

Where $Y$ is the $n$ by $s$ matrix of the observed outputs of all DMUs, $X$ is the $n$ by $m$ matrix of the observed inputs for all DMUs, and $\lambda$ is a $n$-dimensional row vector of weights that forms convex combination of observed DMUs relative to which the subject DMU's efficiency is evaluated. The constraint in this problem simply describe the input requirement set as given by the observed data. Note that a DMU's efficiency is a relative measure. It compares a DMU's performance to the best practice performance implicit in the observed input-output combinations. If different input-output combinations were observed, a DMU's efficiency score would likely change. This idea is the bootstrap performed below.

## 3. The bootstrap

In this method, artificial, or pseudo-samples are drawn from the original data; the statistic is recalculated on the basis of each pseudo-sample; the resulting bootstrapped measures are then used to construct a sampling distribution for the statistic of interest. Note that in order for the bootstrap to work, the empirical distribution of the sample must be a good representation of the underlying population distribution that generated the sample in first place.

We use the efficiency scores calculated from the original data to form pseudo-samples of artificial data. Each artificial data set is similar to the original data set in that both follow the same distributions of inefficiency; this assures that the levels of performance within the bootstrapped results are within the realm of observed behavior.

The efficiency measures being considered in this article are input-based measures; the bootstrap is performed over the original efficiency scores. For this reason only the inputs are adjusted in the formation of the pseudo-samples. The data in the pseudo-samples thus consist of the original output level for all $n$ DMUs, the original input data for the DMU whose efficiency is being calculated, and adjusted input data for the remaining $n-1$ DMUs. After forming a pseudo-sample, the efficiency of a DMU's original input vector is then assessed relative to the technology implicit in it. Recalculating a DMU's efficiency relative to a large number of pseudo-samples generates a sampling distribution for the efficiency score.

To perform our analysis, we modify a form of the bootstrap that is commonly used in the analysis of regression equations. Calculate the technical efficiency of the $t^{th}$ DMU relative to the pseudo-technology implicit in $X_t(b)$ and $Y$ by solving B times the linear program:

$$
\begin{aligned}
min \quad & \theta_t(b) \\
& st: \\
& \lambda.Y \geq y_t \\
& \lambda.X_t(b) \leq \theta_t(b).x_t \\
& \sum_{t=1}^{n} \lambda_t = 1 \\
& \lambda_t \geq 0 \quad t = 1, ..., n.
\end{aligned}
$$

to get the bootstrapped efficiency score $\theta_t^*(b)$ . [8]

After implementing bootstrap method on linear programming, the mean efficiency scores can be obtained differently. And this make possible ranking them.

## 4. An illustration using sections of hospitals in Ardabil

This section discuses the use of DEA to evaluate the efficiency of 9 general sections in Ardabil hospitals. Of great impotance is the choice of inputs and outputs. The following items were considered:

*Inputs* :

1) Number of beds staffed and in operation
2) Number of nuesrs
3) Total operation expenditure
4) Number of physicians
5) Floor area
6) Energy use

⋮

*Outputs* :
1) Number of sicks
2) Quality of care
3) Physicians trined
4) Nurses trained
⋮

In this study first two inputs and outputs are considered. Therefore inputs and outputs data for 9 units DMU's given as follows:

| DMU | INPUT 1 | INPUT 2 | 0UTPUT 1 | OUTPUT 2 |
|---|---|---|---|---|
| 1 | 10 | 2 | 2 | 1 |
| 2 | 3 | 1 | 3 | 4 |
| 3 | 15 | 2 | 12 | 2 |
| 4 | 20 | 4 | 20 | 3 |
| 5 | 17 | 2 | 15 | 3 |
| 6 | 37 | 4 | 30 | 1 |
| 7 | 45 | 4 | 31 | 1 |
| 8 | 9 | 3 | 8 | 4 |
| 9 | 9 | 4 | 4 | 3 |

Table 1. INPUTs and OUTPUTs Data used in the analysis.

The sections of hospitals used two inputs to produce two outputs (Table 1). The results of the equation (1) for these inputs and outputs are summarized in Table 2.

| DMU | Scores |
|---|---|
| 1 | 0.5 |
| 2 | 1 |
| 3 | 0.895 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 0.444 |

Table 2. Original efficiency scores ($\theta_t^*$).

The results do not supply much information to decision makers as it is not possible to distinguish among the performances of many of the sections of hospitals. The bootstrap helps to shed more light upon the performance levels of the observed DMUs. The mean bootstrapped efficiency scores for all DMUs are summarized as follows:

| DMU | Scores | Mean | Ranking |
|---|---|---|---|
| 1 | 0.5 | 0.5021 | 8 |
| 2 | 1 | 0.9965 | 3 |
| 3 | 0.895 | 0.8911 | 7 |
| 4 | 1 | 0.9987 | 2 |
| 5 | 1 | 0.9829 | 6 |
| 6 | 1 | 1 | 1 |
| 7 | 1 | 0.9934 | 4 |
| 8 | 1 | 0.9899 | 5 |
| 9 | 0.444 | 0.4490 | 9 |

Table 3. Mean bootstraped efficiency scores $\theta_t^*(b)$ and Ranking.

Using this method, the efficient scores can be revised. In other words such DMUs lose their efficiency levels, with low perturbation of inputs of DMUs. For example original efficiency scores of $DMU_2$ and $DMU_4$ equal 1. Having an eye on the bootstrapped efficiency scores, it will be obvious that efficiency of $DMU_4$ is more stable than that of $DMU_2$.

## 5. Conclusion

In this paper Data envelopment analysis was used evaluate 9 sections of Hospitals in Ardabil with as inputs number of beds and nurses, and as outputs number of sicks and quality of care. The results of the BCC model for this inputs and outputs are summarized in Table 2. Bootstrap Method is used to rank efficient DMUs .Despite the difficulties and critics that may arise in using the method, it is valuable, because on the other contrary, Methods are able to rank non-extreme DMUs. Using this method, the efficient scores can be revised. The result is shown is Table 3.

## References


[1] Aigner D., Lovell C. A. K. and Schmidt P. (1977). *Formulation and Estimation of Stochastic Frontier Production Function Models*, Journal of Econometrics; 6, 21-37

[2] Asgharian, M., Khodabakhshi, M., and Neralic , L. (2010). *Congestion in stochastic data envelopment analysis: An input relaxation approach.* International Journal of Statistics and Management System; 5(12), 84-106.

[3] Banker R.D., Charnes A., Cooper W.W. (1984). *Some models for estimating tecnical and scale ineffcencies in data envelopment analysis*, Management Science; 30; 9, 1078-1092.

[4] Charnes A., Cooper W.W., Rhodes E. (1978). *Measuring the efficiency of decision making units*, European Journal of Operational Research; 2; 6, 429-444.

[5] Cooper, W. W., Deng, H., Huang, Z., and Li, S. X. (2004). *Chance constrained programming approaches to congestion in stochastic data envelopment analysis*, European Journal of Operational Research; 155(2), 487-501.

[6] Cooper, W. W., Seiford, L. M., and Tone, K. (2000). *Data envelopment analysis. A comprehensive text with models, applications, references and DEA-solversoftware*, Kluwer Academic Publisher.

[7] Ebadi S., Jahanshahloo G.R., Mozeli A.A., Aliev F. (2009) *Determination measure of efficiency using undesirable outputs of DEA.*

[8] Ebadi S., Jahanshahloo G.R., Appl. Comput. Math., V.12, N.1, (2013) *Confidence intervals for dea models efficiency scores by Bootstrapping method D.* News of Baku University, Physico Mathematical sciences series 4:86-91.

[9] Fare R., Grosskopf S. and Lovell, C. A. K. (1985). *The Measurement of Efficiency and Production*, Boston: Kluwer-Nijhoff Publishing.

[10] Farrell M.J. (1957). *The measurement of productive efficiency*, Journal of Royal Statistical Society A; 120; 3, 253-290.

[11] Huang Z., and Li S.X., (1996). *Dominance stochastic models in data envelopment analysis*, European Journal of Operation Research; 95(2), 390-403.

[12] Jess, A., Jongen, H. Th., Neralic , L., and Stein, O. (2001). *Semi-infinite programming model in data envelopment analysis*, Optimization; 49(4), 369-385.

[13] Kall, P. (1976). *Stochastic linear programming*, Berlin: Springer-Verlag.

[14] Khodabakhshi, M. (2009). *Estimating most productive scale size with stochastic data in data envelopment analysis*, Economic Modelling; 26(5), 968-973.

[15] Khodabakhshi, M., and Asgharian, M. (2009). *An input relaxation measure of efficiency in stochastic data envelopment analysis*, Applied Mathematical Modelling; 33, 2010-2023.

[16] Morita Hiroshi and Seiford, L. M. (1999). *Characteristics on stochastic DEA efficiency*, Journal of the Operational Research Society; 42(4), 389-404.

**Said ebadi** -is a Assistant Professor and Faculty member in Mathematics at the Islamic Azad University, Ardabil, Iran. Much of his teaching has been in operations research received his B.S. degree from Tabriz University in 1995, and the M.S degree from Department of Mathematics, Islamic Azad University, Lahijan, Iran.and Ph.D.degree in Applied Mathematics at Azerbaijan Natoninal Academy of sciences University, Baku, Azerbaijan.His major research interests are operations research and data envelopment analysis.