Nigerian Journal of Science and Environment 2025 Volume 23 (1) 1 – 12 ISSN: 3043 – 4440
https://doi.org/10.61448/njse231251

Investigating the discriminating power and efficiency of Data Envelopment Analysis (DEA) model based on Stochastic Data Envelopment Analysis

[1]Ohoriemu, O.B, [2]Okwonu, F.Z. and [3]Mohammad, R.G.
[1]Department of Mathematics, Faculty of Science, University of Delta, Agbor, Nigeria
[2]Department of Mathematics, Faculty of Science, Delta State University, Abraka, Nigeria
[3]Department of Industrial Engineering, Hakim Sabzevari University, Sabzevar, Iran
Corresponding Author's Email: okeoghene.ohoriemu@unidel.edu.ng

Abstract

The problem associated with data envelopment analysis (DEA) is the lack of discrimination power among efficient decision-making units (DMUs), hence yielding many DMUs to be efficient. The issue is highlighted when the number of DMUs evaluated is significantly less than the number of inputs and outputs used in the evaluation. Therefore, other DEA models were proposed in the literature to overcome the discriminant power problems. However, the existing DEA models suffer from some other difficulties such as infeasible solutions for efficient DMUs, the non- uniqueness of the input-output weights, etc. The focus of this article is to propose a stochastic data envelopment analysis (SDEA) model for improving the discriminating power of the DEA and to overcome those drawbacks. In this article, the data set was applied to demonstrate the efficiency of the proposed model. It can be concluded that the proposed SDEA model perform better than other methods in the aspect of discrimination power and efficiency.

Keywords: Data envelopment analysis; Decision making units; Stochastic DEA; Efficiency

Introduction

Data envelopment analysis (DEA) was first proposed by Charnes, *et al.,* (1978) and remained the leading technique for measuring and evaluating the relative efficiencies of a set of homogenous decision-making units (DMUs) based on their respective multiple inputs and outputs. The inputs can consist of labour, materials, energy, machines, and other resources, while the outputs may consist of finished products, services, customer satisfaction, and other forms of outcomes (Ghasemi, *et al.*, 2019). DEA has been the fastest growing discipline in the past three decades covering easily over a thousand papers in the *Operations Research* and *Management Science* discipline. There are large number of DEA applications in environmental performance, especially at the national level (Halická & Trnovská, 2019).

One of the difficulties of the conventional DEA model is the lack of discriminating power among efficient DMUs, hence yielding many of DMUs to be efficient. The issue is highlighted when the number of DMUs evaluated is considerably less than the number of inputs and outputs used in the evaluation. The input and output weights obtained from a DEA analysis may reveal that some inputs or outputs have zero values. This is counter-intuitive particularly in a decision-making exercise, where one expects to utilize all the input-output values that are rated for the DMUs. Hence, it further implies that some of the variables were not used in the evaluation analysis in achieving the final ranking. On the contrary, unrealistic weight distribution and dispersion for DEA also happens when some DMUs are rated as efficient due to exceedingly large weights in a single output and/or exceedingly small weights in a single input. Some techniques such as assurance region (AR) procedure Thompson, *et al.,* (1990), cone ratio envelopment Charnes, *et al.*, (1990) were addressed in DEA literature as strategies to solve problems arising from unrealistic weight distribution and dispersion. However, these methods suffer from some other drawbacks. The AR and cone ratio techniques are highly dependent on the measurement of the input and output units, which may lead to infeasible solutions. In other words, both methods incorporate extra constraints to the model; thus, making it harder to solve the problem.

Subsequently, other DEA models were developed in DEA literature to deal with discrimination power problems, such as the cross-efficiency evaluation technique (Ebrahimi, *et al.*, 2022; Ghasemi, *et al.,* 2019; Sexton, *et al.*, 1986), super-efficiency model (Andersen & Petersen, 1993; Chen, 2005), multiple criteria (or multi-objective) DEA (Li & Reeves, 1999) and stochastic DEA (SDEA) model (Davtalab-Olyaie, *et al.*, 2019; Tsionas, 2021). The super-efficiency DEA model may obtain infeasible solutions for efficient DMUs; particularly, under variable returns to scale (VRS) model. With regards to cross-efficiency evaluation technique, the non-uniqueness of the DEA weights could provide many multiple optimal solutions for DEA models. Other issues such as providing negative input or output weights was also reported by using some of these techniques Ebrahimi, *et al.,* (2022). In Stochastic DEA (SDEA) model, input and/or output data are supposed to be stochastic and it cannot probably be used under deterministic dataset. It can be concluded that the existing methods still possess the following problems: (i) lack of discriminating power among efficient DMUs, hence yielding many DMUs to be efficient, (ii) the negative weight distribution which may reveal that some input or output weights possess negative values, that would not be acceptable in DEA model and (iii) the need of providing proper efficiency values for some DMUs which are not very close or almost equal to zero.

This article is focussed on improving the classical DEA model and existing DEA methods such as goal programming DEA models and bi- objective weighted models in terms of discriminating power. It proposes a suitable DEA model for obtaining positive input and output weights, and improves upon the existing DEA model in terms of providing greater efficiency values for DMUs.

The issue is highlighted when the number of DMUs evaluated is significantly lesser than the number of inputs and outputs used in the evaluation. Therefore, other DEA models were proposed in the literature to overcome the discriminant power problems. However, the existing DEA models suffers from some other difficulties such as infeasible solutions for efficient DMUs, the non- uniqueness of the input-output weights, etc. The focus of this article is to propose a stochastic DEA (SDEA) model for improving the discriminating power in DEA and to overcome those drawbacks. The rest of this article is organized as follow. Section 2 consider different type of DEA models while Section 3 contains data collection, results and discussion. Conclusion follows in Section 4.

The lack of discriminating power among efficient DMUs

The issue is highlighted when the number of DMUs evaluated is considerably less than the number of inputs and outputs used in the evaluation. The input and output weights obtained from a DEA analysis may reveal that some inputs or outputs have zero values. This is counter-intuitive particularly in a decision-making exercise, where one expects to utilize all the input-output values

Nigerian Journal of Science and Environment 2025 Volume 23 (1) 1 – 12 ISSN: 3043 – 4440
https://doi.org/10.61448/njse231251

that are rated for the DMUs. Hence, it further implies that some of the variables were not used in the evaluation judgment in achieving the final ranking. On the contrary, unrealistic weight distribution and dispersion for DEA also happens when some DMUs are rated as efficient due to exceedingly large weights in a single output and/or exceedingly small weights in a single input. Some techniques such as assurance region (AR) procedure Thompson, *et al.*, (1990), cone ratio envelopment Charnes, *et al.*, (1990) were addressed in DEA literature as strategies to solve problems arising from unrealistic weight distribution and dispersion. However, these methods suffer from some other drawbacks. The AR and cone ratio techniques are highly dependent on the measurement of the input and output units, which may lead to infeasible solutions. In other words, both methods incorporate extra constraints to the model; thus, making it harder to solve the problem.

Other DEA models were also addressed in the literature to overcome the discriminating power problems, such as the super-efficiency model (Andersen & Petersen, 1993) and cross-efficiency evaluation technique (Doyle & Green, 1994; Sexton, *et al.*, 1986). However, super-efficiency may obtain infeasible solutions for efficient DMU under evaluation. It can also be mentioned that the super-efficiency does not project on the true efficient frontier but on the fictitious one that results when the DMU being evaluated is removed. The non-uniqueness of the optimal input-output weights in cross-efficiency approach provides many multiple optimal solutions for DEA models.

Data envelopment analysis (DEA) model

Consider the relative efficiency of n DMUs which use m inputs (x_{ij}, $i = 1, \dots, m, j = 1, \dots, n$) to produce s outputs (y_{rj}, $r = 1, \dots, s$, $j = 1, \dots, n$). By assuming that the inputs-outputs data are non-negative and at least one input and one output are positive, we solve the following fractional programming problem for each DMU to obtain the values of input weights (v_i, $i = 1, \dots, m$) and output weights (u_r, $r = 1, \dots, s$) as variables.

$$
\begin{aligned}
& max\ \theta_o = \frac{u_1y_{1o}+u_2y_{2o}+\cdots+u_sy_{so}}{v_1x_{1o}+v_2x_{2o}+\cdots+v_mx_{mo}} \\
& s.t. \\
& \frac{u_1y_{1j}+u_2y_{2j}+\cdots+u_sy_{sj}}{v_1x_{1j}+v_2x_{2j}+\cdots+v_mx_{mj}} \le 1, \quad j = 1, \dots, n, \\
& u_1, u_2, \dots, u_s \ge 0, \\
& v_1, v_2, \dots, v_m \ge 0,
\end{aligned}
\tag{1}
$$

where
y_{rj} = amount of output r associated with DMU_j
u_r = weight associated with output r
x_{ij} = amount of input i associated with DMU_j
v_i = weight associated with intput i.

In DEA model (1), we use the optimal value of the objective function to provide the efficiency value of DMU_o,
which is equal to

$$\theta_o^* = \frac{u_1^*y_{1o}+u_2^*y_{2o}+\cdots+u_s^*y_{so}}{v_1^*x_{1o}+v_2^*x_{2o}+\cdots+v_m^*x_{mo}}.$$

DMU_o is efficient if the optimal value of θ_o (θ_o^*) is equal to one. Otherwise, if $\theta_o^* < 1$, DMU_o is not efficient (i.e., it is inefficient). It should be mentioned that the optimal value of θ_o is always a value between zero and one, $0 < \theta_o^* \le 1$.

The envelopment form of input-oriented DEA model

Nigerian Journal of Science and Environment 2025 Volume 23 (1) 1 – 12 ISSN: 3043 – 4440
https://doi.org/10.61448/njse231251

Consider evaluating the relative efficiency of n DMUs which use m inputs to produce s outputs, in which the m-input-s-output data can be expressed as

$$(x_{ij}, i = 1, \dots, m, j = 1, \dots, n) \text{ and } (y_{rj}, r = 1, \dots, s, j = 1, \dots, n).$$

The dual form of the multiplier input-oriented DEA-CCR model as the envelopment form of input-oriented DEA-CCR model for providing the efficiency value of DMU$_o$ (DMU under evaluation) can be written as follows (Charnes *et al.,* 1978):

$$\begin{aligned} &\theta_o^* = \min \theta \\ &s.t.\ \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io},\ i = 1, \dots, m \\ &\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro},\ r = 1, \dots, s, \\ &\lambda_j \ge 0, j = 1, \dots, n \end{aligned} \qquad (2)$$

where λ_1, λ_2, …, λ_n are non-negative variables in the above model. We use the optimal value of the objective function to figure out the efficiency of DMU$_o$. Thus, θ_o^* (the optimal value of θ_o) is applied to clarify the efficiency value of DMU$_o$, where $0 < \theta_o^* \le 1$. If $\theta_o^* = 1$, we consider DMU$_o$ as efficient; otherwise, it is mentioned as inefficient. The input-oriented DEA-BCC model can be easily obtained by adding the condition $\sum_{j=1}^{n} \lambda_j = 1$ in the model as a constraint.

The envelopment form of output-oriented DEA model

By considering model (1), in the same manner, the output-oriented DEA-CCR model can be written as follows:

$$\begin{aligned} &\phi_o^* = min\ \phi \\ &s.t. \\ &\textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \le x_{io},\ i = 1, \dots, m, \\ &\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \ge \phi y_{ro},\ r = 1, \dots, s, \\ &\lambda_j \ge 0,\ j = 1, \dots, n \end{aligned} \qquad (3)$$

where λ_1, λ_2, …, λ_n are the same as in the model (1). ϕ_o^*, the optimal value of ϕ is greater than or equal to 1. If $\phi_o^* = 1$, DMU$_o$ is efficient; otherwise, it is inefficient.

The proposed stochastic DEA (SDEA) model

Consider, as usual, the input-output observation$\left(x_{ij}\ ,\ \tilde{y}_{rj}\right)$. Assume that the inputs are deterministic while the outputs are random and each output $\tilde{y}_{rj}$ is normally distributed with mean $\mu(\tilde{y}_{rj})$ and variance Var$(\tilde{y}_{rj})$. Further assume that Cov $(\tilde{y}_{rj}\ ,\ \tilde{y}_{rk}\ ) = 0$. The chance-constrained output-oriented DEA-CCR model for DMU$_o$, can be specified as follows (Ray, 2004):

$$\begin{aligned} &\phi_o^* = min\ \phi \\ &s.t. \\ &\textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \le x_{io},\ i = 1, \dots, m, \\ &Pr\left(\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \ge \phi y_{ro}\right) \ge (1 - \alpha),\ r = 1, \dots, s, \\ &\lambda_j \ge 0,\ j = 1, \dots, n \end{aligned} \qquad (4)$$

By setting α at any value between zero and one, that is, we require the inequality restriction involving the outputs to hold with probability 1 – α or higher. Now, the random variable can be defined as

$$u = \sum_{j=1}^{n} \lambda_j y_{rj} - \phi y_{ro}.$$

Then,

$$\mu(u) = E(u) = \sum_{j=1}^{n} \lambda_j \mu(y_{rj}) - \phi \mu(y_{ro}),$$

and

$$Var(u) = \sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(y_{rj}) + (\lambda_o - \phi)^2 Var(y_{ro}),$$

$$\sigma_u = \sqrt{\sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(y_{rj}) + (\lambda_o - \phi)^2 Var(y_{ro})}.$$

Since $y_{rj} (r = 1, \dots, s,\ j = 1, \dots, n)$ is normally distributed, so does the variable u. Hence, the variable

$$z = \frac{u - \mu_u}{\sigma_u}$$

has the standard normal distribution. Thus,

$$Pr\left(\sum_{j=1}^{n} \lambda_j y_{rj} \geq \phi y_{ro}\right) = Pr(u \geq 0)$$

$$Pr\left(z \geq \frac{-\mu_u}{\sigma_u}\right) = Pr\left(z \leq \frac{\mu_u}{\sigma_u}\right) \geq (1 - \alpha).$$

The α value was set at the conventional level of 0.05 by Ray (2004). Therefore, the inequality restriction involving the outputs can be held with probability 95% or higher (see Ray, 2004). Besides, the critical value of the standard normal distribution at the 5% level of significance is 1.96. It can be concluded that

$$\mu_u \geq 1.96\sigma_u$$

and

$$\sum_{j=1}^{n} \lambda_j \mu(y_{rj}) - \phi \mu(y_{ro}) \geq 1.96\sqrt{\sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(y_{rj}) + (\lambda_o - \phi)^2 Var(y_{ro})}.$$

By considering the above inequality, model (4) can be written as follows (Ray, 2004):

$$\begin{aligned} &\phi_o^* = min\ \phi \\ &s.t. \\ &\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io},\ i = 1, \dots, m, \\ &\sum_{j=1}^{n} \lambda_j \mu(y_{rj}) \geq \phi \mu(y_{ro}) + 1.96\sqrt{\sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(y_{rj}) + (\lambda_o - \phi)^2 Var(y_{ro})}, \\ &r = 1, \dots, s, \\ &\lambda_j \geq 0,\ j = 1, \dots, n \end{aligned} \tag{5}$$

By considering models (1) and (5) and assuming that x_{ij} and y_{rj} are random input and output and each input x_{ij} and output y_{rj} is normally distributed with mean $\mu(x_{ij})$ and $\mu(y_{rj})$ and variance $Var(x_{ij})$ and $Var(y_{rj})$. Further assume that Cov(x_{ij} , x_{ik}) = 0 and Cov(y_{rj} , y_{rk}) = 0.

The proposed SDEA model can be considered as follows:

Nigerian Journal of Science and Environment 2025 Volume 23 (1) 1 – 12 ISSN: 3043 – 4440
https://doi.org/10.61448/njse231251

$$
\begin{aligned}
&\theta_o^* = min\, \theta \\
&s.t. \\
&\textstyle\sum_{j=1}^{n} \lambda_j \mu(x_{ij}) - \partial \sqrt{\sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(x_{ij}) + (\lambda_o - \theta)^2 Var(x_{io})} \leq \\
&\theta x_{io}, \;\; i = 1, \ldots, m, \\
&\textstyle\sum_{j=1}^{n} \lambda_j \mu(y_{rj}) + \partial \sqrt{\sum_{j=1, j\neq o}^{n} \lambda_j^2 Var(y_{rj}) + (\lambda_o - 1)^2 Var(y_{ro})} \;\geq \\
&y_{ro}, \;\; r = 1, \ldots, s, \\
&\lambda_j \geq 0, \;\; j = 1, \ldots, n,
\end{aligned}
\tag{6}
$$

where ∂ is the user-specified value, a value between zero and one as the significance level.The advantage of the proposed model (6) can be illustrated with a dataset comprising multiple inputs and multiple outputs of a proper number of DMUs for a period of several days, weeks, months, years, etc. In this case, the mean and standard deviation of each input and output associated with each DMU can be provided and by using the proposed model (6), the efficiency value of each DMU can be achieved.

Results and Discussion

Since banks are important to economic growth as they mobilize savings, provide credit to businesses, and contribute to financial stability and perform important functions for society, the optimal use of resources is necessary to provide better services to customers. In addition, providing the performance and efficiency of different bank branches determines the optimal consumption of resources for each bank branch. Therefore, in field research, a traditional user-cost framework was adopted, with labour (x_1), physical capital (x_2), and purchased funds (x_3) as the inputs, and demand deposits (y_1), short-term loans (y_2), and medium-and-long-term loans (y_3) as the outputs (Kao & Liu, 2019). Table 3.1 shows the means and standard deviations of the six input and output factors for the twenty-five Taiwanese commercial banks in five years, where the monetary terms are expressed in one thousand Taiwan dollars. The input and output factors were assumed to be independent.

Table 1: The means and standard deviations of the input and output factors for the twenty-five Taiwanese commercial banks in five years (Grundfied data set)

Bank	X_1		X_2		X_3		Y_1		Y_2		Y_3	
	Mean	SD	Mean	SD	mean	SD	Mean	SD	Mean	SD	mean	SD
1	6128	48.2	22.5	1.5	646.1	83.42	129	14.5	426.3	45.2	234.7	31

Nigerian Journal of Science and Environment 2025 Volume 23 (1) 1 – 12 ISSN: 3043 – 4440
https://doi.org/10.61448/njse231251

2	6630	129	26.5	1.1	763.1	69.93	146.4	22.5	473.8	43.8	262.7	29.4
3	6209	58.1	21.3	0.47	713	56.12	151.9	21.9	468.1	57.2	233.4	33.6
4	3477	104	11	0.47	351.2	45.16	85.2	17.5	309.1	47.2	114	16.1
5	2074	72.5	6.6	0.56	220.6	20.4	22.4	1.4	119.8	15.1	64.6	22.8
6	2117	67.1	4.1	0.55	218.2	17.98	27.3	3.7	138.6	8.88	55.6	4.95
7	1320	65	3.3	0.09	124.7	24.08	6.2	0.28	54.3	5.51	32.2	0.77
8	1097	173	3.2	0.15	79.1	3.73	4.5	0.4	33.1	3.3	22.4	3.09
9	2144	125	3.6	0.43	175	10.9	20.1	2.8	107.1	12.3	36	34.6
10	4751	737	21.3	4.6	398	60.18	30.4	4.9	279	38.4	133.5	13
11	2294	78.4	5.8	0.19	317	21.65	58.4	29.2	200.6	10.4	138.7	5.48
12	1201	217	4.6	1.22	123.2	13.92	14.9	2.1	99.7	11.1	33	4.95
13	1323	395	1.4	0.46	139.2	22.77	10.7	2.9	116.4	13.3	23.7	22.8

14	1240	190	1	0.31	154.6	32.68	7.4	1.7	118.2	25.4	24.7	5.44
15	5194	193	14.6	1.03	592.2	47.16	96	8.5	351.9	26.1	243.5	9.61
16	1568	379	4.1	0.56	143	19.42	8.7	1.4	105.3	6.64	21.2	21.3
17	1614	199	2.4	0.63	139.1	16.05	10.7	1.2	95.4	14.2	32.4	12.7
18	1508	352	3.8	0.83	155.2	27.7	18.2	3.8	74.4	2.68	59.6	7,9
19	1195	242	3.7	0.89	166	28.18	16	3.2	105.6	10.6	45.7	11.4
20	1540	160	3.5	0.63	172	36.85	15.1	3.1	85.7	8.26	53.6	10.1
21	981	92.8	1.5	0.17	105.2	12.95	7.5	0.88	67.7	5.13	30.8	6.64
22	2409	717	5.2	2.35	168	36.1	11.3	2.9	140.2	16.7	26.9	5.14
23	1230	202	1.8	0.41	104.1	11.76	8.5	2.3	79.7	9.49	36.1	5.06
24	1421	323	3.1	0.25	129.1	16.38	8.2	2.3	78.8	8.78	45.2	4.49
25	1261	218	2	0.86	140.4	27.51	8.3	1.3	91.6	10.2	39.4	6.75

Based on the proposed SDEA model (6) and employing the dataset in Table 1, the result of efficiency is listed in Table 2 and Figure 1 demonstrate the efficiency of the DMUs. The results of efficiency values were provided in two steps:

***Step 1*:** We first provided the results of efficiency of all DMUs (banks) using the proposed SDEA model for $\partial = 0.2$ based on the dataset of twenty-five Taiwanese commercial banks listed in Table 1

***Step 2*:** We further provided the efficiency values of those DMUs (banks) that were efficient in step 1 using the proposed SDEA model for $\partial = 0.45$ (see Table 2).

The results can be interpreted and compared in the following way. From Table 2, DMUs (Banks) 4, 11, 13, 14, and 23 are efficient across $\partial = 0.2$, wherease by $\partial = 0.45$ DMUs (Banks) 4 and 11 are only determined as efficient.

Table 2: Proposed SDEA model results for $\partial = 0.2$ and $\partial = 0.45$ based on dataset in Table 1

Bank	Efficiency values for $\partial = 0.2$		Efficiency values of efficient DMUs in step 1 for $\partial = 0.45$	
	Eff.	Ranking	Eff.	Ranking
1	0.9270	10	—	10
2	0.8864	14	—	14
3	0.9690	6	—	6
4	1	1	1	1
5	0.7305	21	—	21
6	0.8084	20	—	20
7	0.6203	25	—	25
8	0.6469	24	—	24
9	0.7212	23	—	23
10	0.8881	13	—	13
11	1	1	1	1
12	0.9124	11	—	11
13	1	1	0.8710	5
14	1	1	0.9998	3
15	0.9322	8	—	8
16	0.8286	18	—	18
17	0.8270	19	—	19
18	0.8632	16	—	16
19	0.9474	7	—	7
20	0.7284	22	—	22
21	0.8901	12	—	12
22	0.9278	9	—	9
23	1	1	0.9071	4
24	0.8544	17	—	17
25	0.8782	15	—	15

The choice of δ in the SDEA model introduces flexibility in evaluating performance under varying levels of risk preferences or uncertainty assumptions. When δ=0.2, the model highlights the efficiency of decision-making units (DMUs) under minimal variability or operational risks, making it suitable for a best-case scenario analysis. Conversely, when δ=0.45, the model emphasizes robustness and applies more stringent criteria, identifying only those DMUs that consistently maintain efficiency under more challenging conditions. This dual approach allows stakeholders to compare performance outcomes and make decisions aligned with their risk tolerance and confidence in the dataset. Therefore, it can be concluded that for optimistic point of view, the value of ∂ can be specified as 0.2 and for pessimistic point of view, ∂ can be determined as 0.45, Why?. In other words, by increasing the value of ∂, the optimistic point of view changes to the pessimistic point of view. Consequently, there is an opportunity for the decision maker (DM) to decide on which value of ∂ is the best for the scenario under his or her interpretation.

Data envelopment analysis (DEA) is a technique for identifying the best practices of a given set of DMUs whose performance is categorized by multiple performance metrics that are classified as inputs and outputs. When there exist many DMUs as efficient which means the efficiency values of these DMUs are equal to unity, we will not be able to discriminate between rank and these efficient DMUs. In this study, the proposed SDEA model improved the discriminatory power between efficient DMUs by reducing the number of efficient DMUs in second step (see Table 2).

It should be mentioned, it could be possible to reduce the number of efficient DMUs from 5 in first step to 2 in second step by increasing the value of ∂ from 0.2 to .45 which is a good opportunity for the DMUs to decide on which value of ∂ is the best for the scenario under his or her interpretation as shown in Figure 3.1.

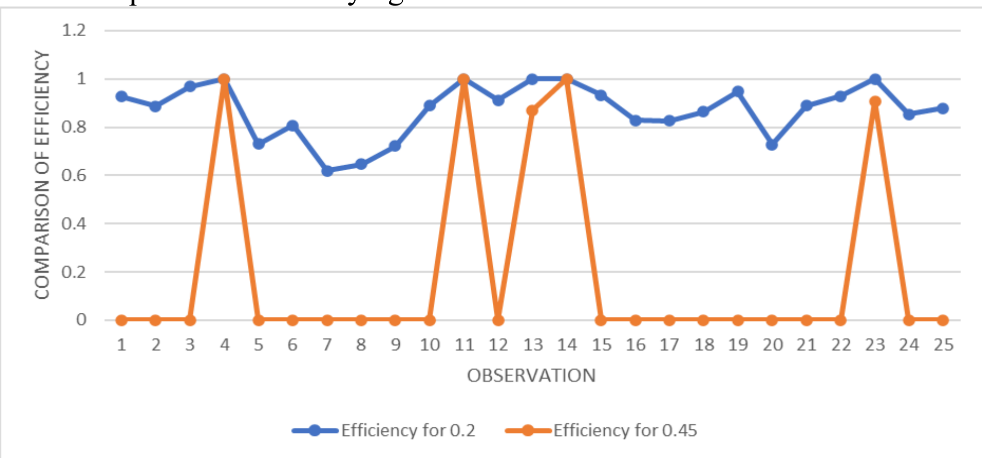

Figure 1: Comparison of efficiency based on ∂ values

Figure 1 compares the efficiency of DMUs under two different ∂ values (0.2 and 0.45). The blue line represents the efficiency scores for ∂ = 0.2, while the orange line represents the efficiency scores for ∂ = 0.45.

The orange line shows sharp peaks for a select few DMUs, indicating instances where the efficiency score is maximized at 1. However, for most DMUs, the orange line remains at or near zero, highlighting a significant reduction in the number of efficient DMUs when ∂ is increased to 0.45.

This sharp contrast demonstrates the improved discriminatory power achieved by the higher ∂ value, which filters out less efficient DMUs, leaving only the most efficient ones.

In contrast, the blue line shows a smoother and more consistent flow, with efficiency values remaining high for many DMUs. This suggests that at ∂ = 0.2, a larger proportion of DMUs are rated as efficient, resulting in reduced discriminatory power between them.

The interplay between these two lines emphasizes the importance of adjusting ∂ values in the SDEA model. Increasing ∂ provides stricter criteria for efficiency, thus reducing the number of efficient DMUs and offering decision-makers better insights into performance differentiation. This balance between inclusiveness (blue line) and stringency (orange line) is critical for tailoring the model to specific scenarios.

Conclusion

It can be seen that the existing methods still possess the issues of the lack of discriminating power among efficient DMUs, hence yielding many DMUs to be efficient and the negative weight distribution which may demonstrate that some input or output weights possess negative values, that would not be acceptable in DEA model and providing the efficiency values for some DMUs which are very close or almost equal to zero.The stochastic DEA (SDEA) model improve the discriminating power in DEA and overcome those drawbacks such as the negative weight distribution and small efficiency values which are very close to zero. This article shows a procedure that can perform better than those reported in DEA literature. In this article, the bank data set was applied to demonstrate the efficiency of the proposed model. It can be concluded that the proposed stochastic data envelopment analysis model in this article perform better than other DEA methods in the aspect of discrimination power.

References

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. Management Science, 39, 1261-1264.

Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1990). Polyhedral cone-ratio models with an illustrative application to large commercial banks Journal of Econometrics, 46, 73-91.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.

Doyle, J., & Green, R. (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. Journal of the operational research society, 45(5), 567-578.

Kao, C., & Liu, S. T. (2019). Stochastic efficiency measures for production units with correlated data. *European Journal of Operational Research*, *273*(1), 278-287.

Ray, S. C. (2004). *Data envelopment analysis: theory and techniques for economics and operations research*. Cambridge university press.

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). The methodology of data envelopment analysis *San Fransisco: Jossey-Bass.*

Thompson, R. G., Langemeier, L. N., Lee, C. T., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics, 46*, 93-108.

Chen, Y. (2005). Measuring super-efficiency in DEA in the presence of infeasibility. *European Journal of Operational Research, 161*, 447-468.

Davtalab-Olyaie, M., Asgharian, M., & Nia, V. P. (2019). Stochastic ranking and dominance in DEA. *International journal of production economics*, *214*, 125-138.

Ebrahimi, B., Dhamotharan, L., Ghasemi, M. R., & Charles, V. (2022). A cross-inefficiency approach based on the deviation variables framework. *Omega*, *111*, 102668.

Ghasemi, M. R., Ignatius, J., & Emrouznejad, A. (2014). A biobjective weighted model for improving the discrimination power in MCDEA. *European Journal of Operational Research, 233*, 640-650.

Ghasemi, M. R., Ignatius, J., & Rezaee, B. (2019). Improving discriminating power in data envelopment models based on deviation variables framework. *European Journal of Operational Research, 278*(2), 442-447.

Halická, M., & Trnovská, M. (2019). National Level, Duality and profit efficiency for the hyperbolic measure model. *European Journal of Operational Research, 278*(2), 410-421.

Li, X.-B., & Reeves, G. R. (1999). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research, 115*, *507-517.*

Tsionas, M. G. (2021). Optimal combinations of stochastic frontier and data envelopment analysis models. *European Journal of Operational Research, 294(*2), 790-800.