MEASURING EFFICIENCY OF MONGOLIAN COMPANIES BY SFA AND DEA METHODS

Batchimeg Bayaraa
Institute affiliation (Finance department, Economics and Business faculty University of Debrecen), Debrecen, Hungary
Bayaraa_batchimeg@yahoo.com

Abstract: *Efficiency measurement usually adopts one of the following analysis, DEA (Data Envelopment Analysis) or SFA (Stochastic Frontier Analysis), but it is not common to use and compare both models in one research. Especially, there is not any research about performance measurement which used Mongolian companies' financial data. The aim of this research is to examine the consistency of efficiency scores from DEA and SFA methods on Mongolian public companies. The financial statements of 100 public companies were obtained from the Mongolian Stock Exchange (MSE) website, from 2012 until 2015. Financial statements were chosen which met the requirements of consistency and accuracy, out of 227 public companies. From initially selected 9 output variables, revenue was chosen as an output variable, while cost of goods sold, operating expenses, and cash are used as input variables based on the stepwise regression result. SPSS (Statistical Package for the Social Sciences) software was used for linear regression to choose the variables; Pearson correlation to examine the correlation between variables and the correlation between efficiency scores of DEA, SFA, and COLS (Corrected Ordinary Least Squares); one-way ANOVA was used to determine statistically significant difference among the methods; and unrelated T-test was used for every pair models. In contrary, Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) were performed in R- Excel statistical program. The average efficiency results indicated that the SFA model exhibited the highest score of 0.75 (TeMode), followed by DEA-VRS (Variable Return to Scale) 49.1 and DEA-CRS (Constant Return to Scale) 33.8. Due to the low-efficiency score, scale efficiency was adopted, and the result showed only 3 companies work in an optimal efficient scale, while 42 companies work below an efficient scale, and 55 companies work above an efficient scale. Unrelated T-test result showed that there was not statistically significant difference among Tej, TeBC, and COLS; TeMode and CRS; CRS and output efficiency.*

Keywords: *Data Envelopment Analysis (DEA); Stochastic Frontier Analysis (SFA); input efficiency; output efficiency; Variable Return to Scale (VRS); Constant Return to Scale (CRS); Corrected Ordinary Least Squares (COLS).*

JEL classification: *C14; C300; L250.*

1. Introduction

Performance evaluation indices are in fact an action guide from what it is towards what it should be (Tehrani et al., 2012). As for corporate performance, it is the measurement for what had been achieved by a company, which is measured by either DEA or SFA, mostly. DEA is a nonparametric method, which has the origin in

production theory as a tool to evaluate production efficiency (Yu, 1994), while SFA is a parametric method which requires production function to evaluate efficiency.
There are a great number of papers which used the DEA method. For example, Fenyves et al., (2015) analyzed applicability of Data Envelopment Analysis (DEA) in the performance measurement, which covered the 5 year period from 2009 to 2014. Tehrani et al., (2012) analyzed the financial performance of 36 companies, and 9 companies were efficient. Zohdi et al., (2012) measured the efficiency of Iranian 12 investment companies' financial statements by CCR and BCC. The empirical results demonstrated that BCC form of DEA was not able to distinguish between efficient and inefficient units. Nikoomaram et al., (2010) researched six-years of data from 24 companies and with the title of "Efficiency Measurement of Enterprises Using the Financial Variables of Performance Assessment and Data Envelopment Analysis".
Similar to the DEA method, the SFA method is also commonly used for measuring efficiency. For instance, Crisci et al., (2016) Technical efficiency with several stochastic frontier analysis models using panel data, Lensink and Meesters (2014) did research about Institutions and Bank Performance: A Stochastic Frontier Analysis, and Price et al., (2017) Production Costs, Inefficiency, and Source Water Quality: A Stochastic Cost Frontier Analysis of Canadian Water Utilities etc.
However, there have been a significant number of studies that performed efficiency measurement on either DEA or SFA, the number of researches which compared both methods is not many. For instance, Ueasin et al., (2015) the technical efficiency of Rice Husk Power Generation in Thailand: Comparing Data Envelopment Analysis and Stochastic Frontier Analysis; Lie-Chien-Lin et al., (2005) application of DEA and SFA on the measurement of operating efficiencies for 27 international container ports; Erkoc (2012) estimation methodology of economic efficiency: Stochastic Frontier Analysis versus Data Envelopment Analysis etc.
As for now, there is not any published research of performance measurement applied DEA and SFA specifically targeted at Mongolian companies.
The objectives of the study were:

- To determine corporate efficiency based on the DEA and SFA.
- To examine the correlation between the efficiency measurement methods.
- To explain the results.

The remainder of this paper is organized as follows: Section two reviews the literature about data envelopment analysis (DEA) and stochastic frontier analysis (SFA). Section three provides data and variables, and the methodology of this study. Section four consists of empirical results and discussion. Finally, conclusions are drawn in section five.

2. Literature Review

For every company, monitoring efficiency is one of the key activities. Efficiency measurement methods can be divided into three main categories: ratio indicators, parametric and nonparametric methods (Vincová, 2005).
A major difference between the parametric and the non-parametric approaches is the estimation method. Whereas the DEA methods rely on the idea of minimal extrapolation, the parametric approaches use classical statistical principles, most notably the *maximum likelihood principle* (Bogetoft and Otto, 2011).

2.1 Data Envelopment Analysis (DEA)

Efficiency is an operation level that produces the greatest amount of output with the lowest amounts of input (Ueasin et al., 2015). DEA is a nonparametric efficiency measuring method which can be determined by either output efficiency or input efficiency. Both input and output efficiency receive a score of 1 for efficient companies. Although input can be the number of hours, the amount of money, financial ratios, time, natural resources, etc. to produce any given output. This study utilizes the following financial data: cost of goods sold, operating expenses and cash for input variables, while revenue is an output variable. There are three different efficiency measures in DEA: technical efficiency *TE*, cost efficiency *CE* and, allocative efficiency *AE*. The relationship between them is easy to derive:

$$CE = AE * TE \text{ (Bogetoft and Otto 2011).}$$

DEA differs by its model supporting scale assumptions: constant return to scale (CRS) and variable return to scale (VRS) (Fenyves et al. 2015). The VRS and CRS models are treated in input oriented forms while the multiplicative model is treated in output oriented form (Banker et al., 2004).

2.2 Stochastic Frontier Analysis (SFA)

SFA is a parametric approach and is suited to measure efficiencies of the stochastic industry for input/output information (Lin and Tseng 2005). Like DEA method, SFA method also receives a score of 1 for efficient companies, although it does not require any efficient company for every observation unless possible inefficiency (u) is equal to zero. In the parametric approach, three main processes have been suggested.

To consider any deviation as noise corresponding to an ordinary regression model.

To consider any deviation as an expression of inefficiency, so-called the deterministic frontier.

Deviations are the results of both noise and inefficiency. This is the stochastic frontier approach (Bogetoft and Otto 2011).

3. Data, variables and research methodology

Financial statements of 100 Mongolian public companies from 2012 until 2015 were downloaded from MSE's website. As for public companies, yearly financial reports are required to be audited, which increases the data reliability. 7 output variables and 13 input variables are initially selected for the research. According to linear regression result (stepwise model), revenue is chosen as an output variable and operating expenses, cost of goods sold and cash were utilized as input variables, which explained the output best (adjusted R square 0.978). Variables which had VIF score above 3 were excluded from research due to multi co-linearity.

As we can see from Table 1, cost of goods sold and revenue are highly correlated (94.9%), while the amount of cash and the amount of operating expenses are loosely correlated (44.4%) with each other.

The descriptive statistics of the inputs and output are shown in Table 2. From the result, we can observe considerable high values of standard deviations among the companies, which indicate that relatively big and small companies are chosen as data.

Table 1: Correlation matrix of the chosen variables

Variables	Revenue	Operating expenses	Cost of goods sold
Revenue	1.000		
Operating expenses	0.822**	1.000	
Cost of goods sold	0.949**	0.654**	1.000
Cash	0.600**	0.444**	0.523**

** Correlation is significant at the 0.01 level (2-tailed).
Source: based on author's own calculation

4. Analysis and results

In the scope of this research, Mongolian 100 public companies' performance measurement was evaluated by DEA and SFA in R- Excel Statistical Program based on their 4 years financial reports from the Mongolian Stock Exchange website.

Table 2: Descriptive statistics of input and output variables

	Minimum	Maximum	Mean	Std. Deviation
Revenue	0.0	255 895 259.8	14 083 072.9	38 462 770.7
Operating expenses	0.0	64 519 117.9	2 635 235.3	7 941 265.8
Cost of goods sold	0.0	191 124 097.6	9 459 880.4	25 434 959.9
Cash	4.0	50 114 496.6	1 091 603.9	4 055 529.7

Source: based on author's own calculation

When the parameter is 0, there is no effect from differences in efficiency, and if it is very large, differences are almost only due to differences in efficiency and not to other kind of uncertainty (Bogetoft and Otto 2011). In table 3, it is seen that the estimated parameter is 1.389, which means that the total error variance is mainly due to inefficiency, whereas the random errors are less significant.

Table 3: Estimation of Cobb-Douglas production frontier using OLS (lm) and SFA

Variables	Coefficient	Std.error	t-value	Pr(>|t|)
(Intercept)	2.591	0.529	4.893	0.000
Cost of goods sold	0.253	0.080	3.140	0.002
Operating expenses	-0.114	0.061	-1.862	0.065
Cash	0.822	0.078	10.456	0.000
Lambda	1.389	0.838	1.656	0.100

Source: based on author's own calculation

From table 4, we see that the percentage of total variation due to variation in efficiency is 65.88%. The estimated variance for the variation in efficiency (u) is

0.314, and is considerably larger than the variation due to random errors (v) 0.162. 65.88% of the total variation is due to inefficiency and the remaining 34.12% is the random variation. The variance for efficiency is larger than the variance of random errors.

Table 4: Estimation variation in SFA

Variance for inefficiency	U	0.314
Variance for random errors	V	0.162
Percentage of inefficiency variation to total variation	100*lambdaˆ2/(1+lambdaˆ2)	65.88%

Source: based on author's own calculation

Table 5 shows the correlation of 3 different models in SFA, which are highly correlated with each other. TeJ and TeBC are the most highly correlated models (99.9%), while other models have also high correlation between 98.2-98.4%. The average efficiency of TeMode model has the highest score of 0.75; however, TeJ model has the lowest score of 0.657 at average efficiency.

Table 5: Correlation among SFA methods and its mean efficiency scores

	TeBC	**TeMode**	**TeJ**
TeBC	1.000		
TeMode	0.982	1.000	
TeJ	0.999	0.984	1.000
Mean efficiency	0.672	0.750	0.653

Source: based on author's own calculation

Figure 1 presents the relationship between revenue and efficiency based on SFA calculation. It is seen that most of the companies' efficiency score is between 0.7-0.8, and the efficiency score collapsed after the score exceeded 0.8.

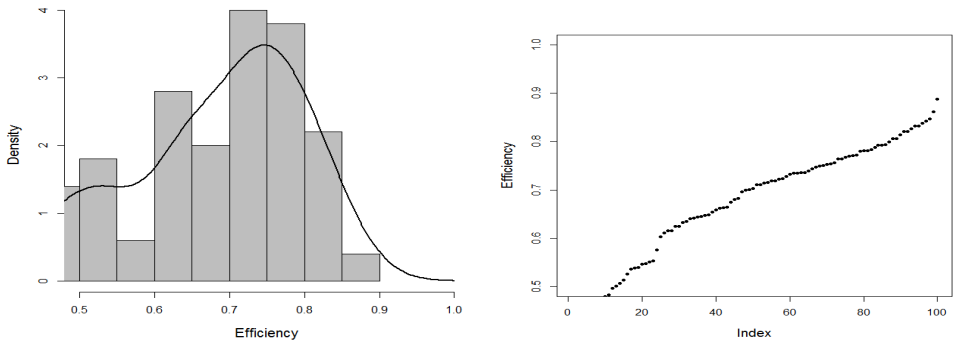

Figure 1: Efficiencies: Histogram, density, and order
Source: based on author's own calculation

Table 6 explains the efficiency results for each method for DEA, SFA, and COLS. The mean efficiency score of SFA is comparatively higher than that of other methods.

As for DEA method, only 12 companies out of 100 are efficient for DEA-VRS, which has the average efficiency score of 49.1%. The number of efficient companies additionally falls in CRS method (average efficiency 33.8%), which shows only 4 companies are efficient, while 70 companies are working between the efficiency score range 0.1-0.4. The number of efficient companies is the highest for DEA-output efficiency (15 companies out of 100); however, average efficiency score is 2.863. It means to get 1 unit of output we have to consume 2.863 units of input in general. In contrary, we can conclude that SFA method, particularly, TeMode gives the highest average efficiency score (75%). TeMode model also has 18 efficient companies, which is the highest number compared with other models. It is noteworthy to mention that SFA method is used to estimate the best technical efficiencies of company, rather than average technical efficiencies of a firm (Lin and Tseng 2005), so that it does not require an efficient company from observation. In contrary, DEA method always has an efficient company for every observation. On the other hand, average efficiency score is the lowest for COLS (28.2%) and under this model, there is only one firm that performed efficiently. In contrast, 99 companies were found in the range of 0.0-0.8.

Table 6: Comparison of DEA, SFA and COLS efficiencies

Efficiency ranges	**Stochastic frontier analysis**			**Data envelopment analysis**			**COLS**
	TeJ	**TeBC**	**TeMode**	**DEA-VRS**	**DEA-CRS**	**DEA-Output**	**TeCols**
0.0-0.3	-	-	-	32	54	21	62
0.3-0.4	4	4	3	14	20	16	21
0.4-0.5	11	8	11	13	8	21	9
0.5-0.6	13	11	9	10	8	14	5
0.6-0.7	27	24	19	9	3	2	1
0.7-0.8	34	39	12	5	1	3	1
0.8-0.9	11	13	16	3	1	6	-
0.9-1.0	-	-	12	2	1	2	-
1.0	-	-	18	12	4	15	1
Minimum	*0.305*	*0.322*	*0.305*	*0.108*	*0,069*	*1.000*	*0.064*
1st quartile	*0.585*	*0.609*	*0.613*	*0.258*	*0,189*	*1.443*	*0.184*
Median	*0.686*	*0.706*	*0.778*	*0.431*	*0,284*	*2.206*	*0.263*
Mean	*0.653*	*0.672*	*0.750*	*0.491*	*0,338*	*2.863*	*0.282*
3rd quartile	*0.752*	*0.767*	*0.940*	*0.636*	*0,416*	*3.374*	*0.352*
Maximum	*0.881*	*0.887*	*1.000*	*1.000*	*1.000*	*14.430*	*1.000*

Source: based on author's own calculation

In Table 7, we see correlation between the efficiency scores of models which are tested by Pearson correlation. Among the methods, SFA models have the highest

correlation with one another between 88.4-100 %. Correlation between SFA (TeMode) model and DEA (output) model has the lowest correlation score (65.2%).

Table 7: Correlation between DEA, SFA and COLS efficiencies

Models	Stochastic frontier analysis			Data envelopment analysis			COLS
	Tej	TeBC	TeMode	VRS	CRS	Output	TeCOLS
Tej	1						
TeBC	1.000	1					
TeMode	0.889	0.884	1				
VRS	0.686	0.678	0.664	1			
CRS	0.748	0.738	0.706	0.793	1		
Output	0.700	0.693	0.652	0.847	0.724	1	
TeCOLS	0.873	0.862	0.791	0.713	0.838	0.689	1

Correlation is significant at the 0.01 level (2-tailed).
Source: based on author's own calculation

From figure 2, it is clear that there are several firms with a DEA efficiency score of 1, which have much lower COLS efficiency. There is even a firm with a DEA—Input efficiency-VRS, CRS and output efficiency of 1.0 which has COLS efficiency score of 0.3. Moreover, there are efficient companies for DEA and COLS, but not for the SFA method. It can be explained that DEA method always has at least 1 efficient DMU (Decision Making Unit), while SFA does not require any fully efficient DMU unless variance of inefficiency is equal to 0.

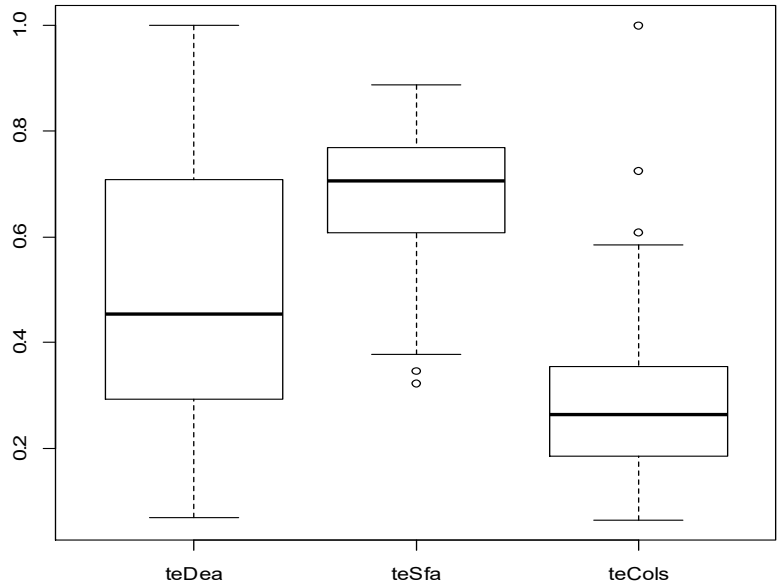

Figure 2: Comparison DEA (teDea) and SFA (teSFA) efficiencies, and COLS (teCols)
Source: based on author's own calculation

Scale efficiency in DEA: To measure the loss from not operating at optimal scale size, we use *scale efficiency SE*, which is determined by the ratio of input efficiency in a CRS model to that in a VRS model, i.e.
SE=E(CRS)/E(VRS) (Bogetoft and Otto 2011).
As we can see from Figure 3, only 3 companies (3%) are working in an efficient scale, while 97% of the companies are working either above (55%) or below (42%) the efficient scale.

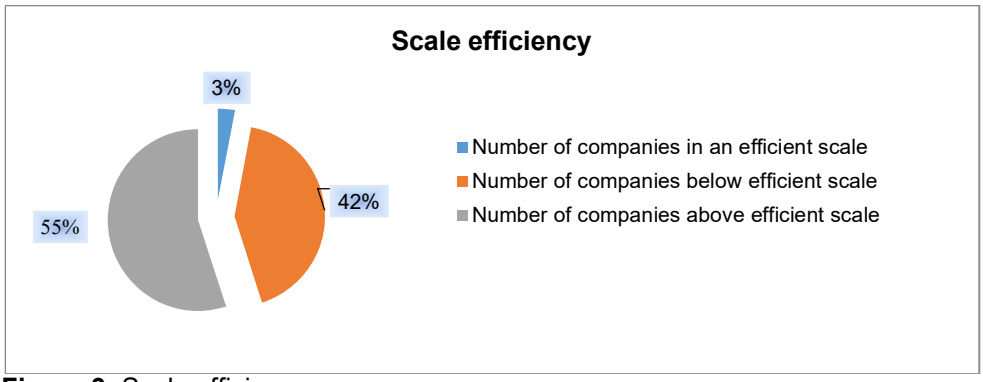

Figure 3: Scale efficiency
Source: based on author's own calculation

According to table 8, it is seen that three different models of SFA have similar efficiency results such as: minimum efficiency scores are between 0.305-0.322; however, only TeMode method results an efficient company with the efficiency score of 1. Maximum efficiency score of Tej and TeBC are 0.881 and 0.887, respectively. DEA efficiency score shows the super efficiency for input and output efficiency which is not restricted to either above or below one. Maximum efficiency score of VRS is 5.175, which means if we use 1 unit of input, we can produce 5.175 units of output. Efficiency score for DEA-output efficiency is opposite of the other models which prefers the lowest score. To clarify, the highest efficiency boundary of DEA-output is 14.434 shows to get 1 unit output 14.434 units of inputs are required, while the lowest efficiency boundary is 1 which explains the efficient company.

Table 8: Descriptive statistics of efficiencies by DEA (super efficiency), SFA and COLS

Methods	Models	Minimum	Maximum	Mean	Std. Deviation
SFA	Tej	0.305	0.881	0.653	0.130
	TeBC	0.322	0.887	0.673	0.126
	TeMode	0.305	1.000	0.757	0.197
DEA	VRS	0.000	5.175	0.582	0.716
	CRS	0.069	2.793	0.369	0.357
	Output	1.000	14.434	1.915	3.610
COLS	TeCols	0.064	1.000	0.282	0.151

Source: based on author's own calculation

Table 9 details companies which are efficient in scale efficiency have the highest efficiency scores for every model.

Table 9: Comparison of the most efficient companies by DEA, SFA, and COLS

Stochastic frontier analysis			Data envelopment analysis			COLS	Scale efficiency
teJ	teBC	Temode	Input-VRS	Input CRS	Output		
0,756	0,771	0,953	4,842	1,805	0,445	0,363	Yes
0,881*	0,887*	1,000*	1,011	2,793*	0,357*	1,000*	Yes
0,716	0,734	0,844	5,175*	1,461	Inf	0,300	Yes

**The most efficient companies*

Source: based on author's own calculation

Table 10 highlights if there is a statistical significant difference between the models. From the result, it is seen that there is not statistically significant difference among Tej, TeBC, and COLS; TeMode and CRS; CRS and output efficiency.

Table 10: Statistical significance in efficiency among the methods

	Tej	TeBC	TeMode	VRS	CRS	Output	COLS
Tej	1						
TeBC	equal	1					
TeMode	unequal	unequal	1				
VRS	unequal	unequal	unequal	1			
CRS	unequal	unequal	Equal	unequal	1		
Output	unequal	unequal	unequal	unequal	equal	1	
COLS	equal	Equal	unequal	unequal	unequal	unequal	1

Source: based on author's own calculation

5. Conclusion

This paper aimed to measure the efficiencies of Mongolian 100 public companies from 2012 until 2015, by DEA, SFA, and COLS. Moreover, I tried to compare the results of the models and find out if any statistically significant difference exists. The paper uses linear regression to choose the variables, the benchmarking package of R statistical program to evaluate the efficiencies, one-way ANOVA, and unrelated T-test to check the statistically significant difference among the models. Revenue is chosen as an output variable, while cost of goods sold, operating expenses and cash are used as input variables, according to stepwise regression result.

The analysis shows the mean efficiency score of SFA (TeMode) model is the highest (0.757), followed by SFA (TeBC-0.673), SFA (TeJ-0.653), DEA (VRS-0.582), and DEA (CRS-0.369), while the lowest score is COLS 0.282. In SFA (TeMode) model, 18 companies are efficient which is the highest number; in contrary, COLS shows only 1 company works efficiently, and SFA (TeJ) and SFA (TeBC) do not reveal any efficient companies. Unrelated T-test results show that efficiencies are not significantly different among Tej, TeBC, and COLS; TeMode and CRS; CRS and output efficiency. However, other models significantly differ by their efficiency results from one another.

References

1. Banker, R.D. et al., 2004. Returns to scale in different DEA models. *European Journal of Operational Research*, 154(2), pp.345–362.
2. Bogetoft, P. & Otto, L., 2011. Benchmarking with DEA, SFA and R. *International Series in Operations Research & Management Science Series Editor :*,
3. Crisci, A., Siviero, L. & D'Ambra, L., 2016. Technical efficiency with several stochastic frontier analysis models using panel data. *Electronic Journal of Applied*, 9(4), pp.736–759.
4. Erkoc, T., 2012. Estimation Methodology of Economic Efficiency: Stochastic Frontier Analysis vs Data Envelopment Analysis. *International Journal of Academic Research in Economics and Management Sciences*, 1(1), pp.1–23.
5. Fenyves, V., Tarnoczi, T. & Zsido, K., 2015. Financial Performance Evaluation of Agricultural Enterprises with DEA Method. *Procedia Economics and Finance*, 32(15), pp.423–431.
6. Lensink, R. & Meesters, A., 2014. Institutions and bank performance: A stochastic frontier analysis. *Oxford Bulletin of Economics and Statistics*, 76(1), pp.67–92.
7. Lin, L.C. & Tseng, L.A., 2005. Application of DEA and SFA on the measurement of operating efficiencies for 27 international container ports. *Proceedings of the Eastern Asia Society for Transportation Studies, Vol 5*, 5, pp.592–607.
8. Nikoomaram, H., Mohammadi, M. & Mahmoodi, M., 2010. Efficiency Measurement of Enterprises Using the Financial Variables of Performance Assessment and Data Envelopment Analysis. *Applied Mathematical Sciences*, 4(37), pp.1843–1854.
9. Price, J.I. et al., 2017. Production Costs, Inefficiency, and Source Water Quality: A Stochastic Cost Frontier Analysis of Canadian Water Utilities. *Land Economics*, 93(1), pp.1–11.
10. Tehrani, R., Mehragan, M.R. & Golkani, M.R., 2012. A Model for Evaluating Financial Performance of Companies by Data Envelopment Analysis- A Case Study of 36 Corporations Affiliated with a Private Organization. *International Business Research*, 5(8), pp.8–16.
11. Ueasin, N., Liao, S.-Y. & Wongchai, A., 2015. The Technical Efficiency of Rice Husk Power Generation in Thailand: Comparing Data Envelopment Analysis and Stochastic Frontier Analysis. *Energy Procedia*, 75, pp.2757–2763.
12. Vincová, I.K., 2005. Using Dea Models To Measure ... Using Dea Models To Measure Efficiency 1. *Biatec*, (1), pp.24–28.
13. Zohdi, M. et al., 2012. Data envelopment analysis (DEA) based performance evaluation system for investment companies: Case study of Tehran Stock Exchange. *African Journal of Business Management*, 6(16), pp.5573–5577.