Journal of Data Science 69-78 , DOI: 10.6339/JDS.201801_16(1).0004

Efficiency Analysis of Manufacturing Firms Using Data Envelopment Analysis Technique

Tagdira Naznin Smriti 1 ,Md Hasinur Rahaman Khan1

Abstract: Efficiency analysis is very useful and important to measure the performance of the firms in com- petitive market of rapidly developing country like Bangladesh. The more efficient firms, and the decision making units (DMUs) are usually referred as benchmarking units for the development. In this study, efficiency scores are obtained using the non-parametric Data Envelopment Anal- ysis (DEA) technique for 1007 manufacturing firms in Bangladesh from the enterprise survey data. The DEA is used to calculate weights for inputs and outputs by assigning the maximum efficiency score for a DMU under evaluation. Total 29 firms are found efficient under variable returns to scale assumption. The significant determinants behind the inefficiency found in this analysis include mainly the firm size, manager's experience in respective sector, annual losses due to power outage, number of production workers.

Keywords: Efficiency analysis; Variable returns to scale; Competitive market; Enterprise survey

1 Introduction

Efficiency Analysis is a challenging issue for measuring the performance of various sectors. Per- formance measurement and benchmarking is very important for measuring efficiency. Frontier efficiency techniques can be used for finding the efficiency score which is related to the 'best practices'. Efficiency of the firms also make substantial changes in our economy [5]. An ac- commodating business environment is one that encourages firms to operate efficiency [10]. For measuring the efficiency the surveys are repeated overtime to track changes and benchmark the effects of reforms on firms performance. The Enterprise survey covers several topics of the business environment as well as performance measure for each firm.

Two frontier techniques are usually used for capturing the efficiency of the firms. They are Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) [19]. The standard DEA efficiency scores are sensitive to measurement error, specially with small samples. It is a non-parametric technique used in the estimation of production functions and has been used extensively to estimate measures of technical efficiency in a range of industries [8], [16]. If there is no knowledge about the population or parameter but still it is required to test the hypothesis of the population then non parametric approach is used. If the information about the population is completely known by means of its parameters then parametric approach is used. The primary focus of these approach is modeling the production or performance function of DMUs. Since we do not have clear idea about the population of Bangladeshi industry's efficiency hence we decided to implement the non parametric DEA approach. Like the stochastic production frontiers, DEA estimates the maximum potential output for a given set of inputs, and has primarily been used in the estimation of efficiency. However, again like the AHP and ANP approaches, DEA also can be used to estimate capacity utilization.

Since in seminal work of Farrell in 1957, several empirical studies have been conducted on firm efficiency. Efficiency can be decomposed into technical and allocatable [9]. Farrells method extended by relating the restrictive assumption of constant return to scale and of strong disposability of inputs [11]. Banker and Morey (1986) adapted DEA to allow an analysis of efficiency on the basis of exogenous and non-exogenous fixed inputs and outputs. The major demerit of this approach is its reliance on benchmark ratios [24]. These benchmarks could be arbitrary and may mislead an analyst. Financial ratios don't capture the long-term performance, and aggregate many aspects of performance such as operations, marketing and financing [23]. The total factor productivity (TFP) growth rate in Malaysia analysed and decomposed the total factor productivity into technological change and technical efficiency change [15]. How does DEA perform– there are many works that help to get insights including [20] that analyzed

the efficiency of the Olympic sports with the data analytics approach, [14] that analyzed the optimism in the automotive industry and [18] etc.
The non-parametric approach has been used since recent past years. The research has suggested that the kind of mathematical programming procedure used by DEA for efficient frontier estimation is comparatively robust [22]. A key advantage of DEA over other approaches previously examined is that it more easily accommodates both multiple inputs and multiple outputs. DEA is a linear programming model to measure efficiency under the assumption of constant returns to scale [6] and extended by Banker et al. (1984) to allow variable returns to scale [3]. Grosskopf provided a good survey of statistical inference in non-parametric linear programming frontier models, also analyzing the asymptotic properties of the estimators[12]. A large number of papers have been extended and applied the DEA methodology [7], [21], [1], [2], [14], [18], [4] etc.
In this study one of our contributions is that we have extended the existing key performance indicators with a new, powerful benchmarking tool that addresses the limitations of the indicators currently used in the firms. We have estimated the efficiency score in selected firms in Bangladesh using DEA for output oriented VRS. The determinants of efficiency score can be obtained by DEA approach.

2 Methodology

Modern frontier efficiency methodologies, like more traditional techniques such as financial ratio analysis aims at benchmarking firms of an industry against each other. Frontier efficiency techniques measure a company's performance relative to the "best practice" of the most efficient companies in the same industry. Efficiency estimates are standardized between 0 and 1, with the value 1(0) assigned to the most (least) efficient firm. Using a production frontiers to compute technical efficiency is the simplest and most wide-spread approach in the frontier efficiency literature, generally by means of one of two fundamental orientation. The input orientation aims at minimizing inputs conditional on given output levels. The output orientation, on the other hand, maximizes output levels conditional on a given input consumption.
Two types of DEA model in the light of orientation are usually used. These are – Input Oriented DEA model and Output Oriented DEA model. "By how much can inputs be reduced while maintaining the same level of output?". This model focuses on the measurement of the variations in input use and an increase in efficiency will be achieved by reducing inputs proportionally holding outputs constant. On the contrary the corresponding output oriented question could be, "By how much can output be increased while keeping the level of inputs constant?". Output oriented DEA model focuses on the measurement of variation in output produced and an increases in efficiency will be achieved by increasing outputs proportionally holding the input quantities constant. According to return to scale there are two types of DEA model. These are CRS (Constant Returns to Scale) and VRS (Variable Returns to Scale) model. The CRS assumption is appropriate only for DMU's decision making units that are operating at an optimal scale. In a situation where DMU's are not operating at an optimal scale the technical efficiency measure will be mixed with scale efficiency. Hence, to separate the two inefficiency scores it is opted for applying variable returns to scale.
The best way to introduce DEA is via the ratio form. For each DMU we would like to obtain a measure of the ratio of all outputs over all inputs, such as $ujyi/vjxi$, where u is an $M\times 1$ vector of output weights and v is a $k\times 1$ vector of input weights. To select the optimal weights we specify the mathematical programming problem

$$\max_{u,v}(ujyi/vjxi),$$

such that

$$ujyi/vjxi \leq 1 \;, j = 1, 2, \ldots, N, \qquad u, v \geq 0.$$

This involves finding values for u and v, such that the efficiency measure of the i-th DMU is maximized, subject to constraint that all efficiency measures must be less than or equal to one. One problem with this particular ratio formulation is that it has an infinite number of solutions. To avoid this one can impose the constraint vjxi = 1 which provides

$$\max \mu,v(\mu jyi)$$

such that

$$vjxi = 1,$$

$$\mu jyj - vjxj \leq 0 \text{ , } j = 1, 2, \ldots, N, \qquad \mu, v \geq 0.$$

Here the notation change from u and v to μ and v reflects the transformation. The form is known as the multiplier form of the linear programming problem. Using the duality in linear programming, one can derive an equivalent envelopment form of this problem

$$min_{\theta,\beta}\theta$$

such that

$$-yi + Y\beta \geq 0. \; \theta xi - X\beta \geq 0, \beta \geq 0.$$

Here θ is a scalar and β is a N×1 vector of constraints. The value of θ obtained the efficiency score for the i-th DMU. It will satisfy $\theta \leq 1$, with a value of 1 indicating a point on the frontier and hence a technically efficient DMU. There are various types of software for calculating efficiency based on DEA approach namely, DEAP, Warwick DEA, IDEAS. The DEAP software (version 2.1) is used for data analysis. This identifies efficient projected points which have input and output mixes which are as similar as possible to those inefficient points, and that it is also invariant to units of measurement.

3 Data and Variables

The data used in this study is taken from Enterprise Survey (funded by the World Bank) website and downloaded freely. The World Bank interviewed a representative sample of the private sector in 4 of the most active economic regions in Bangladesh. The regions are—Dhaka, Chittagong-Sylhet, Khulna, Rajshahi. The sample considered of 1442 business establishment, surveyed from April 2013 through September 2013 and it is a cross-sectional data. As the data consists of 1442 observations, with different types of firm– manufacturing firm, retail firm and services (non retail firm). We conduct our study for manufacturing firm and for this the data consists of 1007 observations. Each manufacturing firm is considered as a single decision making unit (DMU).

For conducting the study we mainly considered five variables which are the total annual sales(y), total annual cost (x1), total annual cost of raw materials(x2), total annual cost of electricity(x3) and permanent full time workers(x4). Total annual sales is considered in this study as output variable which to be maximized and the other four variables are considered as input variables which are our available resources and we want to make best use of these. Here we work with the logarithm transformation of the variables which make the value of the variables smaller and interpretable.

4 Analysis and Results

For measuring the efficiency we consider the variables y, x1, x2, x3, x4 which are total annual sales, total annual cost, total annual cost for raw materials, total annual cost for electricity and permanent full time workers respectively. The unit for y, x1, x2, x3 are Bangladeshi currency (BDT) and the unit for x4 is the number of worker. The summary statistics are given in the following Table1.

Table 1: Summary Statistics for the Variables

Variable	Obs	Mean	Standard Deviation	Minimum	Maximum
y	1007	4.55×10^{08}	6.34×10^{09}	2×10^{05}	2×10^{11}
x_1	1007	2.77×10^{07}	1.19×10^{08}	4×10^{04}	2.32×10^{09}
x_2	1007	1.58×10^{08}	5.35×10^{08}	6×10^{03}	7×10^{09}
x_3	1007	4.45×10^{07}	2.86×10^{07}	1.2×10^{03}	6×10^{08}
x_4	1007	2.46×10^{03}	6.76×10^{02}	4	1×10^{4}

DEA assigns weights to the inputs and outputs of a DMU that give it the best possible efficiency. Table2shows the descriptive statistics for efficiency using the DEA approach. From Table2we can see the mean efficiency under constant returns (CRS) to scale assumption is 0.835 and for variable returns to scale (VRS) assumption is 0.877, using DEA. So the technical efficiency is slightly greater for VRS than CRS. Basically, VRS model is essentially the CRS with an additional constraints added to the linear programming problem.

Table 2: Descriptive Statistics for Efficiency Using DEA

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum
CRS	1007	0.835	.0485	0.757	1
VRS	1007	0.877	.0431	0.77	1
SCALE	1007	0.952	.0281	0.86	1

Under the variable return to scale assumption, the technical efficiency scores (=1) along with peer counts for the efficient firms are obtained in Table3. We get total 29 efficient manufacturing firms. The others firm are inefficient because their efficiency score are less than 1. Technical efficient DMUs are peer of themselves only. Scale efficiency is defined as the region in which there are constant returns to scale in the relationship between outputs and inputs and irs for increasing returns to scale and DRS for decreasing returns to scale. Now five inefficient DMU and their corresponding information is given in Table4.

Table 3: Some Efficient DMUs with Their Information

DMUs	Peer Count	Technical Efficiency	Scale Efficiency
192	786	1	1(-)
339	608	1	0.957(drs)
702	488	1	0.919(drs)
292	464	1	0.959(drs)
135	343	1	1(-)

These inefficient DMUs is expressed as linear combinations of the efficient DMUs for efficiency perspective. Suppose for DMU 5 we get

$$\boldsymbol{Y}_5 = \boldsymbol{\lambda}_{339}\boldsymbol{Y}_{339} + \boldsymbol{\lambda}_{706}\boldsymbol{Y}_{706} + \boldsymbol{\lambda}_{702}\boldsymbol{Y}_{702},$$

Table 4: Some Inefficient DMUs with Their Information

DMUs	Peer Count	Technical Efficiency	Scale Efficiency	Peers	Peer weights
5	0	0.844	0.884(drs)	339,706,702	0.200,0.180,0.680
19	0	0.928	0.898(drs)	339,706,292	0.240,0.228,0.152,0.380
27	0	0.855	0.938(drs)	702,339,352	0.217,0.653,0.130
490	0	0.893	0.920(drs)	135,292,192,889	0.222,0.371,0.30,0.376
517	0	0.996	0.884(drs)	158,702,337	0.257,0.207,0.536

where Y denotes the decision making unit and λ stands for the peer weight. From the equation we can see that the DMU 5 is a linear combination of DMU 339, 706 and 702 and these three DMUs are the efficient DMUs. Peer weights determine what weights should be assigned for the respective peers. For this equation, the peers are 339, 706, 702 and the peer weights assigned to them are 0.200, 0.180, 0.680 respectively. The technical efficiency for DMU 5 is 0.844. That is DMU 5 should be able to increase the inputs by 15.6% without increasing any input, the input level would remain same. Similar interpretation holds for the others.

Table 5: Efficiency Scores of Firms of Different Size Using DEA Approach

Firm Size	Mean	Standard Error	95% Confidence Interval	
Small	0.8748	.0025	0.8698	0.8798
Medium	0.8734	.0022	0.8691	0.8777
Large	0.8850	.0022	0.8806	0.8894

The efficiency scores for different firm size are given in Table5. Small sized firm have employees >= 5 and <= 20 whereas medium sized firm have employees >= 20 and <= 99 and large sized firm have employees >= 100. We see that the mean efficiency for small firm is 87.48%, for medium firm is 87.34% and for large sample is 88.50% implying some little better efficiency for large firm. The efficiency scores of the manufacturing firms depend on the other factors like managers experience of the firm, labour cost, production workers, etc.

For efficiency measurement by DEA method it can be identifiable that which variables are responsible for efficiency. If we run the regression for efficiency score obtained by DEA then we get the determinants. The plausible determinants (variables) are selected after checking the significance of the individual variables. And here we use the variables a6a (firm size), b3 (percentage held by largest owners), b2a (private domestic individuals, companies or organization), b2b (private foreign individuals, companies or organization), b6 (full-time employees at start-up), b7 (manager's experience), c9b (annual losses due to power outages), d1a3 (percentage of total sales of the main product), d3a (national sales), l3a (production workers) etc as possible determinant. Now we run the regression for finding out the more efficient variable and find out the p-value. From the p-value we can take the decision about the determinants. The variables and the p-values is given in Table6.

From the p-value we can say that the variable a6a (Firm size for medium) is effective as determinant for efficiency at the 6% level of significance. Similarly b7 (Manager's experience in sector) is significantly effective for achieving efficiency in the firm and at the 1% level of signif- icance. And c9b (Annual losses due to power outages), d3a (National sales), l3a (Production workers) are also useful for getting efficient score above at the 10% level of significance. Hence the fitted regression model is

$$\hat{y} = 0.75 - 0.007a6a + 0.00063b7 - 0.00013d3a - 0.00001l3a + 3.64 \times 10^{-10}c9b,$$

Table 6: Estimates of the Parameters of the regression model

Variable	Coefficient	Standard Error	P-value	95% confidence interval	
a6a	-.0071	.00376	0.059	.00026	-.01453
b3	-.00008	.00006	0.149	-.00020	.00003
b2a	.0012	.00086	0.154	-.00046	.00293
b2b	.0013	.00087	0.122	-.00036	.00306
b6	.00002	.00002	0.294	-.00001	.00006
b7	.00063	.00015	<0.001	.00033	.00094
c9b	3.64×10^{-10}	1.97×10^{-10}	0.065	-2.30×10^{-11}	7.52×10^{-10}
d1a3	.00008	.00006	0.193	-.00004	.00021
d3a	-.00013	.00005	0.016	-.00023	-.00002
l3a	-.00001	5.81×10^{-06}	0.065	-.00002	6.88×10^{-07}
cons	0.7533	.087823	<0.001	0.58095	0.92583

where yˆ denotes estimated efficiency score for output variable returns to score obtained by data envelopment analysis.

5 Conclusions

Early research on performance measurement was based on the traditional financial ratio anal- ysis. However, several authors have suggested that frontier efficiency provides a new, powerful performance measurement technique and a valuable addition to them existing performance measures in the field. Efficiency techniques might be helpful in overcoming the ambiguities of traditional financial ratios, as they summarize different characteristics of the firm in a single and easy to interpret performance indicator.

For calculating the efficiency score the non-parametric DEA approach is used. The results also illustrate the diversity of different firms in terms of the performance and emphasize the relevance of benchmarking in identifying the “best practice”. Actually efficiency techniques measure a firm’s performance relative to the most efficient firms. Frontier efficiency analysis is used here to evaluate the performance of manufacturing firms. In this paper, efficiency of manufacturing firms surveyed under the enterprise survey is obtained by Data Envelopment Analysis. We also obtain the factors, most importantly we can mention firm size and manager’s experience which have highly influence on the efficiency scores of the manufacturing firms in Bangladesh as expected.

References

[1] Ahmadi, V., & Ahmadi, A. (2012). Application of data envelopment analysis in manufac- turing industries of Iran. Interdisciplinary Journal of Contemporary Research in Business, 4(8), 534–544.

[2] Akgöbek, Ö & Yakut, E (2014). Efficiency measurement in Turkish manufacturing sector using Data Envelopment Analysis (DEA) and Artificial Neural Networks (ANN). Journal of Economic & Financial Studies, 2(3), 35–45.

[3] Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficienci Data Envelopment Analysis. Management Science, 30(9), 1078–1092.

[4] Barouni, M., Amiresmaieli, M. R., Shahravan, A. & Amini, S. (2017). The Efficiency Assessment of Dental Units Using Data Envelopment Analysis Approach: The Case of Iran. Iran J Public Health, 46(4), 552–559.

[5] Baten, M. A. Kamil, A. A., & Fatama, K. (2009). Technical Efficiency in Stochastic Frontier Model: an Application to the Manufacturing Industry in Bangladesh. Australian Journal of Basic and Applied Sciences, 3(2), 1160–1169.

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

[7] Coelli, T. (1996). A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, CEPA Working Paper No. 96/07. Centre for Efficiency and Productivity Analysis, University of New England, Armidale.

[8] Constantin, P. D., Martin, D. L., & Y Rivera, E. B. B. R. (2009), Cobb-Douglas, Translog Stochastic Production Function and Data Envelopment Analysis in Total Factor Produc- tivity in Brazilian Agribusiness. The Flagship Research Journal of International Conference of the Production and Operations Management Society, 2(2), 20–34.

[9] Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society A, 120, 253–281.

[10] Fernandes, A. M. (2008), Firm Productivity in Bangladesh Manufacturing Industries, World Development, 36(10), 1725–1744.

[11] Färe, R., Grosskopf, S., & Lovell, CA K. (1985). The measurement of efficiency of produc- tion. Springer Science & Business Media.

[12] Grosskopf, (1996). Statistical Inference and non parametric efficiency: a selective survey .
Journal of Productivity Analysis, 7, 161–176.

[13] Hettler, B., Sorokina, N., Tanai, Y. & Booth, D. (2015). Analysis optimism in the automa- tive industry: A post-ballout boost and methodological insights. Journal of Data Science, 13, 473–494.

[14] Mujaddad, A. G. & Ahmad, H. K. (2016). Measuring Efficiency of Manufacturing Industries Pakistan: An Application of DEA Double Bootstrap Technique. Pakistan Economic and Social Review, 54 (2), 363–384.

[15] Jajri, I.(2007). Determinants of total factor productivity growth in Malayesis. Journal of economic Cooperation, 28(3), 41–58.

[16] Khoshroo, A., Mulwa, R., Emrouznejad, A. & Arabi, B. (2013). A non-parametric Data Envelopment Analysis approach for improving energy efficiency of grape production. En- ergy, 63, 189–194.

[17] Li, Z., Liu, S. & Tian, M. (2014). Collective Behavior of Equity Returns and Market Volatility. Journal of Data Science 12, 545–562.

[18] Limaei, S. M. (2013). Efficiency of Iranian forest industry based on DEA models. Journal of Forestry Research, 24(4), 759–765.

[19] Moreira, V. H., & Bravo-Ureta, B. E. (2010), Technical efficiency and metatechnology ratios for dairy farms in three southern cone countries: a stochastic meta-frontier model, Journal of Productivity Analysis, 33(1), 33–45.

[20] Salmon, J. L. & Harrison, W. K. (2017). A Data Analytics Approach to Evaluation of Competition in the 2012 Summer Olympics. Journal of Data Science, 15, 373–390.

[21] Simar, L., & Wilson, P. W. (2007). Estimation and inference in two stage, semi-parametric models of production processes. Journal of Econometrics, 136(1), 31–64.

[22] Seiford, L. M., & Thrall, R. M. (1990). Recent developments in DEA: the mathematical programming approach to frontier analysis. Journal of econometrics, 46(1), 7–38.

[23] Sherman, H. D., & Gold, F. (1985). Bank branch operating efficiency: Evaluation with data envelopment analysis. Journal of Banking & Finance, 9(2), 297–315.

[24] Yeh, Q. (1996). The Application of Data Envelopment Analysis in Conjunction with Fi- nancial Ratios for Bank Performance Evaluation. Journal of Operational Research Society, 47, 980–988.