

A Hybrid Approach of DEA and Multivariate Regression Analysis for Performance Measurement of Nationalized Indian Banks

Bhagat Kishanrao Gayval* and Vishnudas Bajaj

*Dept. of Statistics,
Dr BAM University Aurangabad, Maharashtra, India*

*Corresponding author e-mail: gayvalbk@gmail.com

ABSTRACT

Objective: The main objective of this paper is developing new hybrid model of CCR model and BCC model of DEA as non-parametric technique with Multiple Regression Analysis as parametric technique to ranking Nationalised Banks of India. We have verified ranking results obtained through three methods with Statistical test as non-parametric test.

Research Methods: In this paper we have applied Data Envelopment techniques such as input oriented Constant Returns to Scale model and input oriented Variable Returns to Scale model. These models integrated with Multiple Regression Analysis technique for ranking Indian Banks. Also these obtained results have been verified using Spearman's Rank Correlation test.

Results: In this study, 20 Nationalised banks are selected to identify the ranking of each banks using above mentioned three methodologies. The results indicate that the Bank of Baroda, Punjab National Bank and IDBI Bank Ltd are efficient frontier banks out of 20 banks using DEA models. As per methodology of Statistical Multivariate Regression Analysis, the Bank of Baroda, IDBI Bank Ltd and Punjab National Bank were getting first, second and third rank respectively.

Conclusion: We have concluded that the CCR model, BCC model and Multiple Regression Analysis techniques are positively correlated with each other as per Spearman's rank correlation test result.

Keywords: Linear Programming Problem, Data Envelopment Analysis, Multivariate Regression Analysis, Indian Banking.

INTRODUCTION

Indian banking sector is one of the largest banking system in the world. Indian banking system is mainly divided into Public Sector Banks and Private Sector

Banks. The Public Sector Banks are further divided into two categories as Nationalised Banks and SBI group banks and Private Sector Banks are divided into Foreign Banks

and Domestic Private Banks. The Nationalised Banks in India play the pivotal role of Indian economy. The major market has been captured through Nationalised Banks. Our research has major focus on to improve efficiencies and effectiveness of Nationalised Banks of India. The purpose of this paper is to propose a hybrid methodology based on Data Envelopment Analysis (DEA) and Statistical Regression analysis that addresses to this issue of efficiency using data from Nationalised Indian banks.

Over the past two decades, data envelopment analysis (DEA) has become a popular methodology for evaluating the relative efficiencies of decision making units (DMUs), many researches has been chosen banks are decision making units. They have proved the different DEA models are useful for efficiency measurement of banks and its target setting. Data Envelopment Analysis is an approach to estimate the production function of organizations and organizational units and enables the assessment of their efficiency. The comparison of empirical results produced by DEA and Multivariate Regression Analysis is still uncommon in bank efficiency literature. Multivariate Regression Analysis is useful tool to predict the independent value of variable from one and more than one dependent variable. Regression analysis is a statistical tool that is used to investigate the relationship between variables. It is a parametric technique that requires the specification of a production function.

DEA provides an estimation of the production function to which each individual Decision Making Unit's (DMU's) efficiency score can be compared. Furthermore, DEA offers a number of advantages over traditional techniques including its ability to identify reference units for each DMU. These characteristics

prove to be a very useful managerial tool as it aids in establishing potential causes and methods of improvement for the identified inefficient.

Literature review

Data Envelopment Analysis (DEA) is a linear programming technique to assess the efficiencies of decision making units (DMU) when multiple variables are taken into account. Literature related to efficiency studies can be traced back to Farrell⁴, who treated the production frontier as the basis for efficiency assessment. The Charnes, Cooper and Rhodes (CCR)¹ described a mathematical programming formulation for the empirical evaluation of relative efficiency of a Decision Making Unit (DMU) on the basis of the observed quantities of inputs and outputs for a group of similar referent DMUs. The Banker, Charnes and Cooper (BCC)⁶ provided a formal link between DEA and estimation of efficient production frontiers via constructs employed in production economics. Satye¹⁰ used DEA to study the relative efficiency of Indian banks in the late 1990's with that of banks operating in other countries. He found that the public sector banks have a higher mean efficiency score as compared to the private sector banks in India, but found mixed results when comparing public sector banks and foreign commercial banks in India. Kumbhakar and Sarkar¹¹ estimated the cost efficiency of public and private sector banks in India by using the stochastic cost frontier model with specification of translog cost function.

The Seiford and Zhu⁸ examined the performance of the top 55 US banks using a two-stage DEA approach. Results indicated that relatively large banks exhibit better performance on profitability, whereas smaller banks tend to perform better with respect to marketability. Drake and Howcroft¹² assessed the relative efficiency

of UK clearing bank branches using DEA method. This paper utilized the basic efficiency indices and extended the analysis by examining the relationship between size and efficiency. Many of these studies find that state-owned banks are more efficient than private and foreign banks (Bhattacharyya and Pal¹⁹; Sharma²⁰), find that foreign banks are actually the most efficient. R. Kiani Mavi¹⁷ estimated the ranking of banks branches using DEA and multivariate regression analysis, results indicate that the efficiency calculated from CCR model and multivariate regression model has positively correlated. Premachandra²¹ examine the capability of DEA in assessing corporate bankruptcy by comparing it with logistic regression (LR). The main objective of this paper to assesses efficiency of Indian banks using DEA technique and Multivariate Regression Analysis techniques.

MATERIAL AND METHODS

Data Envelopment Analysis

Data has been collected from RBI website for evaluation. Data Envelopment Analysis is a linear programming procedure for a frontier analysis of inputs and outputs. The input-oriented DEA model under the assumption of constant return to scale & variable return to scale can be used for calculation of input-oriented technical efficiency. In this study, we estimate the CCR model proposed by Charnes et al. (1978) and BCC model proposed by Banker et al. (1984), which allows for variable returns to scale. The input-oriented CRS & VRS model can be written as:

The basic input-oriented CRS model for 'n' DMUs with 'm' inputs and 's' outputs proposed by CCR, the relative efficiency score of pth DMUs is given by

$$Max Z_p = \frac{\sum_{k=1}^s V_k Y_{kp}}{\sum_{j=1}^m U_j X_{jp}}$$

$$s.t. \quad \frac{\sum_{k=1}^s V_k Y_{ki}}{\sum_{j=1}^m U_j X_{ji}} \leq 1 \quad \forall i$$

$$V_k, U_j \geq 0 \forall k, j$$

Where, k = 1 to s (number of outputs);

j = 1 to m (number of inputs);

i = 1 to n (number of DMUs);

Y_{ki} = amount of output k produced by DMU i;

X_{ji} = amount of input j utilized by DMU i;

V_k = weight given to output k and U_j = weight given to input j.

The fractional programme shown in Equation (1) can be reduced to LPP as follows:

$$Min \quad \sum_{j=1}^m U_j X_{jp}$$

$$s.t. \quad \sum_{k=1}^s V_k Y_{kp} = 1$$

$$\sum_{k=1}^s V_k Y_{ki} - \sum_{j=1}^m U_j X_{ji} \leq 0 \quad \forall i$$

$$V_k, U_j \geq 0 \quad \forall k, j$$

This model is called CCR input-oriented DEA model. The efficiency score of 'n' DMUs is obtained by running the above LPP 'n' times.

Multivariate Regression Analysis

The Multivariate linear regression model can be return in below form

$$Y = a + b_i x_i + \varepsilon \quad (3)$$

Where Y is the dependent variable and $x_i ; i = 1, 2, 3 \dots n$ is the i^{th} independent variable. Therefore $b_i ; i = 1, 2, \dots n$ is the coefficient of x_i , a is constant and ε is the estimation error term.

RESULTS AND DISCUSSION

In this paper, we have analyzed efficiency of Indian Banks using hybrid approach of DEA and Regression analysis. In this paper, we have taken data of 20 Nationalised Indian Banks during financial year of 2012-13. We have used two input measure such as interest expended (X1), operating expenses (Y2) and one output measure as total income (Y) (total income is summation of Interest income and other income). Data has been collected from RBI website for the financial year 2012-13.

In this paper, we have used the input-orientation CCR model, input-orientation BCC model of DEA and Multivariate Regression analysis model to evaluate performance measurement of Nationalised Banks and also comparative study on parametric and non-parametric techniques. We have applied Spearman's rank correlation test to testing the result obtained from CCR, BCC and Multivariate regression.

The table 2 shows efficiency and ranking of each bank using input oriented CCR model and input oriented BCC model, last column show ranking obtained through Multiple regression analysis. The both CCR and BCC result indicates that Bank of Baroda, Punjab National Bank and IDBI Bank Ltd. are efficient frontier and their ranks are one. Other than these three banks, efficient frontier banks using BCC model are Bank of Maharashtra, Canara Bank, Corporation Bank, Dena Bank, Punjab and Sind Bank, and United Bank of India. The

remaining banks are inefficient and their scores are less than one. The CCR and BCC efficiency score of DMUs are calculated by using R Software and detail result has shown in CCR & BCC columns of Table 2.

We have used multivariate regression model for this data, then the linear regression equation as:

$$Y = -5688 + 1.11X_1 + 1.964X_2 + \varepsilon$$

Error of estimation can be computed by subtracting Y from the actual amount of the output variable (income) for each DMU. DMUs are ranked based on descending error. The results are shown in last two columns of Table 2. The Multiple regression result indicates that the Bank of Baroda is topper bank out of these banks and its rank is one. The IDBI Bank Ltd and Punjab National Bank are getting second and third rank respectively.

The below descriptive Statistics summary show overall Nationalised Banks efficiency obtained through input oriented Constant Return to Scale model of DEA. Result indicates that out of 20 banks only three banks are efficient and having score one.

Descriptive Statistics Summary of Overall Banks using CCR model using R software-

```
>summary(e)
```

Summary of efficiencies

The technology is crs and input orientated efficiency

Number of firms with efficiency==1 are 3

Mean efficiency: 0.957

```
---
```

Eff range	#	%
0.8 <= E < 0.9	3	15
0.9 <= E < 1	14	70
E == 1	3	15

Min.1st Qu.Median Mean 3rd Qu. Max.
0.8865 0.9412 0.9632 0.9572 0.9785 1.0000

The below descriptive Statistics summary show overall Nationalised Banks efficiency obtained through input oriented Variable Return to Scale model of DEA. Result indicates that out of 20 banks, nine banks are efficient (i.e.45% banks distribution data covered) and having score one.

Descriptive Statistics Summary of Overall Banks using BCC model using R software-
>summary(e)

Summary of efficiencies

The technology is vrs and input orientated efficiency

Number of firms with efficiency==1 are 9

Mean efficiency: 0.98

Eff range	#	%
0.8<= E <0.9	1	5
0.9<= E <1	10	50
E ==1	9	45

Min.1st Qu.Median Mean 3rd Qu. Max.
0.8952 0.9654 0.9917 0.9803 1.0000 1.0000

The below Spearman's rank correlation test result show, it has applied for testing comparison between rank obtained through CCR model, BCC model and Multiple Regression model. Result indicates that CCR model and Multivariate regression model are positively correlated with each other and its correlation is 0.961.Also result indicates that Multiple regression and BCC model of DEA are positively correlated with each other but as compare to CCR model are less correlated.

See table no. 3

CONCLUSION

This paper presents a unique approach of DEA &Multivariate Regression Analysis for evaluating Indian banks performance. We have considered group of Nationalised Indian Banks (i.e. 20 banks), it has been analyzed for effectiveness using DEA & Multivariate Regression Analysis. We have investigated the performance of Indian banks by using input orientated CCR, BCC based DEA model &Multivariate Regression Analysis (Result has been shown in table 2). The result shown that the Bank of Baroda, IDBI banks ltd and Punjab National Banks has most efficient. We have tested and verify the relative efficiency results of CCR model, BCC model and Multivariate regression model using non-parametric Spearman rank test. The results show that there is positively and significantly correlated between the set of rankings. This study provides scope for further research using larger input sample size to test the robustness of the results.

REFERENCE

1. Charnes, A., Cooper, W.W. and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operations Research*, 2: 429-444.
2. Collier, T., Johnson, A.L. and Ruggiero, J. (2011) 'Technical efficiency estimation with multiple inputs and multiple outputs using regression analysis', *European Journal of Operational Research*, Vol. 208, No. 2, pp.153-160.
3. Aigner, D.J., C.A.K. Love11 and P. Schmidt, 1977, Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21-37.

4. Farrel, M.J., 1957, the measurement of productive efficiency, *Journal of the Royal Statistical Society A* 70, 253-281.
5. Cooper, W.W., Seiford, L.M. and Tone, K. (2000) *Data envelopment Analysis*, Boston: Kluwer.
6. Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30, 1078–1092.
7. Wang, N. S., Yi, R. H., & Wang, W. (2008). Evaluating the performances of decision making units based on interval efficiencies. *Journal of Computational and Applied Mathematics*, 216, 328–343.
8. Seiford, L., & Zhu, J. (1999). Profitability and marketability of the top 55 US commercial banks. *Management Science*, 45, 1270–1288.
9. Sengupta, J. (1995). *Dynamics of data envelopment analysis: Theory of systems efficiency*. Boston, MA: Kluwer Academic Publishers.
10. Satye, M, (2003). Efficiency of Banks in developing Economy. The case of India. *European Journal of Operational Research*. 148(3). 662-671.
11. Kumbhakar. S.C. and Sarkar, S. (2004). *Deregulation, Ownership and Efficiency in Indian Banking: An application of Stochastic Frontier Analysis*. IGIDD working paper. Available at: www.igidr.ac.in/conf/finwrk/workshop.pdf.
12. Drake, L., & Howcroft, B. (2002). An insight into the size efficiency of a UK bank branch network. *Managerial Finance*, 28, 24–36.
13. Sun, S., & Lu, W. (2005). Evaluating the performance of the Taiwanese hotel industry using a weight slacks-based measure. *Asia-Pacific Journal of Operational Research*, 22, 487–512.
14. Bogetoft, Peter and Lars Otto (2013), *Benchmarking with DEA and SFA*, R package version 0.23.
15. Coelli, T. J., D. S. P. Rao, C. J. O'Donnell and G. E. Battese (2005). *An Introduction to Efficiency and Productivity Analysis*. Second ed. New York: Springer.
16. Kumbhakar, S. C., S. Ghosh and J. T. McGuckin (1991). "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms." *Journal of Business and Economic Statistics* 9, 279-286.
17. R.Kiani Mavi et al. (2015), Ranking banks branches using DEA and multivariate regression models. *Int. J. Operational Research*, 24, 245-261.
18. Ataullah, A. and Le, H. (2008) 'Economic reforms and bank efficiency in developing countries: the case of the Indian banking industry', *Applied Financial Economics*, Vol. 16, No. 9, pp.653–663.
19. Bhattacharyya and S. Pal. (2013). Financial reforms and technical efficiency in Indian commercial banking: A generalized stochastic frontier analysis. *Review of Financial Economics*, 22(3):109–117.
20. S.Sharma, D. Raina, and S. Singh. (2012) Measurement of technical efficiency and its sources: An experience of Indian banking sector. *Journal of Economics and Management*, 6: 35–57.
21. Premachandra, I.M., Bhabra, G.S. and Sueyoshi, T. (2009) 'DEA as a tool for bankruptcy assessment: a comparative study with logistic regression technique', *European Journal of Operational Research*, Vol. 193, No. 2, pp.412–424.
22. E. Thanassoulis. (1993). *A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments*,

- the Journal of the Operations Research Society, Vol. 44, No. 11, pp. 1129-1144.
23. Kooreman P. (1994). Data envelopment analysis and parametric frontier estimation: Complementary tools, *Journal of Health Economics*, 13: 345-46.
 24. Cooper, W.W., Tone, K., 1997. Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. *European Journal of Operational Research* 99, 72-88.
 25. Friedman, L., Sinuany-Stern, Z., 1997. Scaling units via the canonical correlation analysis and the data envelopment analysis. *European Journal of Operational Research* 100 (3), 629-637.
 26. Sinuany-Stern, Z., Friedman, L., 1998. Data envelopment analysis and the discriminant analysis of ratios for ranking units. *European Journal of Operational Research* 111, 470-478.
 27. Azadeh, A., Seraj, O., Asadzadeh, S.M. and Saberi, M. (2012). An integrated fuzzy regression-data envelopment analysis algorithm for optimum oil consumption estimation with ambiguous data', *Applied Soft Computing*, Vol. 12, No. 8, pp.2614-2630.
 28. Banker, R., Gadh, V.M. and Gorr, W. (1993). A Monte Carlo comparison of two production frontier estimation methods: corrected ordinary least squares and data envelopment analysis', *European Journal of Operational Research*, Vol. 67, No. 3, pp.332-343.
 29. Coelli, T. and Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: with application to European railways', *European Journal of Operational Research*, Vol. 117, No. 2, pp.326-339.
 30. Collier, T., Johnson, A.L. and Ruggiero, J. (2011). Technical efficiency estimation with multiple inputs and multiple outputs using regression analysis', *European Journal of Operational Research*, Vol. 208, No. 2, pp.153-160.
 31. Ifeagwu E.N., Ogu. C.J. and Alagbu E. (2015). Enhancing the performance of wideband code division multiple access (WCDMA) network using antenna diversity technique, *American Journal of Computer Science and Information Technology*, Vol.3, No.1, pp.014-25.
 32. Das, A. and Ghosh, S. (2009). Financial deregulation and profit efficiency: a nonparametric analysis of Indian banks', *Journal of Economics and Business*, Vol. 61, No. 6, pp.509-528.
 33. Premachandra, I.M., Bhabra, G.S. and Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: a comparative study with logistic regression technique', *European Journal of Operational Research*, Vol. 193, No. 2, pp.412-424.
 34. Wang, Y.M., Luo, Y. and Lan, Y.X. (2011) 'Common weights for fully ranking decision making units by regression analysis', *Expert Systems with Applications*, Vol. 38, No. 8, pp.9122-9128.

Table 1: Data table for financial year 2012-13

Sr. No.	DMU's	Interest expended	Operating expenses	Total Income
		Input-X1	Input -X2	Output-Y
1	Allahabad Bank	125693	29581	189126
2	Andhra Bank	91527	20372	139571
3	Bank of Baroda	238814	59467	388273
4	Bank of India	228849	53315	356749
5	Bank of Maharashtra	65801	17966	105254
6	Canara Bank	261989	51420	372309
7	Central Bank of India	161231	42323	235280
8	Corporation Bank	119082	19968	169420
9	Dena Bank	65163	12997	95549
10	Indian Bank	93684	27509	151805
11	Indian Overseas Bank	154248	34078	226496
12	Oriental Bank of Commerce	130036	26652	193595
13	Punjab and Sind Bank	56991	11193	77573
14	Punjab National Bank	270368	81651	461092
15	Syndicate Bank	116666	31788	182951
16	UCO Bank	121702	21766	177039
17	Union Bank of India	175819	45122	276767
18	United Bank of India	67642	15039	103181
19	Vijaya Bank	71739	13630	96589
20	IDBI Bank Ltd.	196912	31344	282838

Table 2: Ranking using CCR, BCC and Multivariate Regression models

DMU's	CCR		BCC		Multivariate Regression	
	Efficiency	Rank	Efficiency	Rank	Error	rank
Allahabad Bank	0.942	13	0.953	10	-2802	15
Andhra Bank	0.971	6	0.989	3	3653	8
Bank of Baroda	1.000	1	1.000	1	12084	1
Bank of India	0.979	3	0.980	6	3704	7
Bank of Maharashtra	0.963	9	1.000	1	2618	10
Canara Bank	0.938	14	1.000	1	-13800	19
Central Bank of India	0.887	18	0.895	12	-21121	20
Corporation Bank	0.978	4	1.000	1	3710	6
Dena Bank	0.964	8	1.000	1	3380	9
Indian Bank	0.957	11	0.988	4	-524	12
Indian Overseas Bank	0.937	15	0.944	11	-5960	18
Oriental Bank of Commerce	0.971	7	0.980	5	2599	11
Punjab and Sind Bank	0.899	16	1.000	1	-1982	14
Punjab National Bank	1.000	1	1.000	1	6309	3
Syndicate Bank	0.944	12	0.960	8	-3292	16
UCO Bank	0.984	2	0.994	2	4889	4
Union Bank of India	0.961	10	0.967	7	-1324	13
United Bank of India	0.972	5	1.000	1	4250	5
Vijaya Bank	0.897	17	0.957	9	-4123	17
IDBI Bank Ltd.	1.000	1	1.000	1	8394	2

Table 3: Spearman's Rank correlation test

Sperman's Rank Correlations					
			CCR	BCC	MLR
Spearman's rho	CCR	Correlation Coefficient	1.000	.562**	.961**
		Sig. (2-tailed)	.	.010	.000
		N	20	20	20
	BCC	Correlation Coefficient	.562**	1.000	.639**
		Sig. (2-tailed)	.010	.	.002
		N	20	20	20
	MLR	Correlation Coefficient	.961**	.639**	1.000
		Sig. (2-tailed)	.000	.002	.
		N	20	20	20
**. Correlation is significant at the 0.01 level (2-tailed).					