An Output Oriented Window Analysis Using Two-stage DEA in Heart Hospitals

Mirpouya Mirmozaffari [1], Mostafa Zandieh [2] and Seyed Mojtaba Hejazi [3]

[1,3]Msc. Student, Department of Industrial Engineering, Faculty of Industrial and Mechanical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran (Email:m.mirmozaffari@gmail.com)

[2] Associate Professor, Department of Industrial Management and accounting, Faculty Shahid Beheshti University, G.C. Tehran, Iran (Email: m_zandieh@sbu.ac.ir)

Abstract: *Between all lethal disease, heart attacks illnesses are reflected as the most prevalent. Data envelopment analysis (DEA) delivers a relative efficiency measure for peer decision-making units (DMUs) with multiple inputs and outputs. Evaluating of hospitals is one of the major applications in DEA. In this study, the output oriented is used. The window analysis model was presented as a model that is utilized to guess the Technical Efficiency (TE), pure technical efficiency (PTE), and Scale Efficiency (SE). A new-two stage DEA model is considered for efficiency evaluation of DMUs. The data covers a six-year span from 2011 to 2016 for 12 local heart hospitals. Three inputs (which is two inputs for the first stage and one for the intermediate element of a two-stage model) and two outputs are chosen and these factors reflect the main function of hospitals. Finally, actual, target and potential improvement of the smallest and largest TE, PTE, and SE average of the first hospital are compared. The system was implemented in Banxia Frontier Analyst software.*

Keywords: *data envelopment analysis, window analysis, intermediate element, two-stage model, frontier software*

1. Introduction

Medical specialists conduct different studies on heart diseases and collect material of heart patients, their indications and disease progression. Gradually are described patients with common diseases who have typical symptoms. Yearly 17.3 million people roughly die from heart disease. Thus, there is valued information hidden in their dataset to be mined [1].

The greatest communal used technique, Data Envelopment Analysis (DEA) based on linear programming, draws a frontier of best practices and shows which health establishments are efficient or inefficient. DEA gives the magnitude of inefficiency and specifies the means of refining efficiency by giving objectives projections for each of the inputs and/or outputs exclusively [2].

To find the unidentified trends in heart disease, all the available hospital's datasets are applied to an exclusive model and their DMUs efficiency are compared. A dataset for 12 hospitals with two inputs, one intermediate element and two outputs in a novel two-stage method for DEA window analysis which are CCR_{OO} and BCC_{OO} to check and validate the changes between hospitals are used. After running the DEA model in Banxia Frontier Analyst software with every unit in the data, this study categorizes all units into three groups: High efficient (H-efficient) with 100% efficiency scores, efficient with between 90% and 100% efficiency scores, and inefficient with less than 90% efficiency scores.

2. Background and literature review

Global increasing number of patients have encouraged scientists to do wide-ranging research to disclose concealed patterns in medical datasets. This section provides a synopsis of previous computational studies on pattern acknowledgement in hospitals. Finally, the gap in existing literature, which was the foremost motivation of this study is also provided. Caballer-Tarazona et al (2010) measured the efficiency of 22 hospitals in the Valencian portion of Spain in expressions of three units [3]. Therefore, pattern recognition in hospitals can be addressed through different computational models. Other respected works, focused on diverse computational aspects of

hospitals on different models including data mining can be mentioned [4-5-6-7]. Finally, the output-oriented model is used. A unique utilizing window analysis and a novel two-stage model is applied in this study.

3. Dataset Description

The standard dataset, assembled in this study covers 6 periods (2011-2016), which is collected from 12 hospitals, under the administration of National Health Ministry. The Number of Doctors (NN) and the Number of Beds (NB) in the hospitals are inputs in the first stage. The Numbers of Nurses and secretaries (NN) is an intermediate element. The Outpatient Treated (OT) and the Inpatient Treated (IT) in the hospitals are outputs in the second stage. So, two inputs, one intermediate element and two outputs for the first hospital in all periods (2011-2016) are presented in Table I.

TABLE I. The inputs, intermediate elements, and outputs for 6 DMUs in the first hospitals

Period	First input	Second input	Intermediate elements	First output	Second output
	NN (units)	ND (units)	NB (units)	OT (units)	IT (units)
2011	16200	200	1450	25300	5693
2012	11046	158	6100	17200	4463
2013	15000	260	560	24100	5000
2014	13601	188	7100	16100	1800
2015	14300	201	2450	17693	4500
2016	12100	303	482	27500	9800

4. Research methodology

4.1. The CCR and BCC model

1) *CCR Model:* The CCR models reflects a stationary or constant return to scale (CRS), which means that a comparative increase in all inputs results in the same proportional increase in outputs. The efficiency of a given DMU is calculated using the CCR model as follows:

$$Min \ \theta_p$$

$$s.t \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_p x_{ip} \qquad ,i = 1, \dots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{rp} \qquad ,r = 1, \dots, s \qquad (1)$$

$$\lambda_j \geq 0, \qquad j = 1, \dots, n$$

Where θ_p represents the technical efficiency score of unit DMU, λ_j signifies the dual variables that recognize the standards for inefficient units. If θ_p is set to one, then the surveyed DMU is considered technically efficient and lies on the efficiency frontier that is composed from the set of efficient units. DEA measures the efficiency of each reflection relative to the frontier that covers all the explanations. Inefficient DMUs can be improved (moved to the efficient frontier) with suggested directions for development which are the points along the frontier. The distance to the efficiency frontier offers a measure of the efficiency.

1) *BCC Model:* Conversely, the BCC model by Banker-Charnes-Cooper altered the Constant Return to Scale (CRS) notion to Variable Return to Scale (VRS). The DMU operates under variable returns to scale and it is assumed that an increase in inputs does not result in a proportional change in the outputs. The BCC model divides the Technical Efficiency (TE) derived from the CCR model into two parts:

- *Pure Technical Efficiency (PTE):* PTE, which ignores the influence of scale size by only comparing a DMU to a unit of comparable scale and measures how a DMU utilizes its sources.
- *Scale Efficiency (SE):* SE, which measures how the scale size affects efficiency. If after applying both CRS, VRS model on the same data, there is a metamorphosis in the two technical efficiencies, this shows that DMU has a scale efficiency and can be calculated by:

$$SE = TE/PTE \tag{2}$$

The BCC is characterized as follows:

$$Min\ \theta_p$$

$$s.t \sum_{j=1}^{n} \lambda_j\, x_{ij} \leq \theta_p x_{ip} \qquad , i = 1, \dots, m$$

$$\sum_{j=1}^{n} \lambda_j\, y_{rj} \geq y_{rp} \qquad , r = 1, \dots, s \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0,\ j=1,\dots,n$$

4.2. A new approach in DEA two stage model

In recent years, Kao et al., 2008 [8]; Wang et al., 2010 [9] and Hosseinzadeh lotfi et al., 2012 [10], proposed various solutions regarding the two-stage model. A new two-stage model has been proposed in this study:

In the proposed solution the two-stage model is considered as a single stage one, where the intermediate elements are considered as part of final inputs. Consider the input, intermediate element, and output are denoted by X, Z, and Y, then the proposed model treats X+Z, and Y as input and output respectively. This model is suggested to determine the efficiency of a two-stage process and prevent the dependency to different weights. With respect to aforementioned information, the proposed model is denoted in Figure 1.
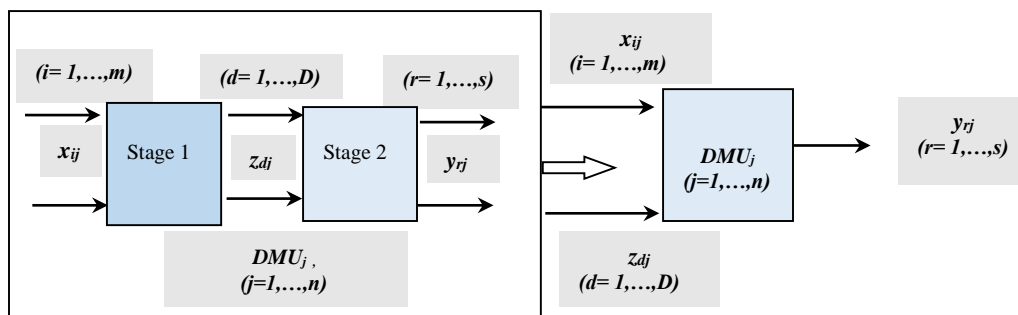

Fig. 1: Conversion of the two-stage model to one-stage model

Figure 1 particularizes the proposed model. In fact, NN with ND units represent X or inputs, NB unit is Z or intermediate elements and OT with IT units are Y or outputs, which were presented in table 1. DEA, as a decision analysis tool, is vast in literature since it does not concentration on discovery a general connection for all units under assessment in the sample. DEA authorities every unit in the data to have its own production function and then it evaluates the efficiency of that single unit by comparing it to the efficiency of the other units in the dataset. Finally, in a more detailed discussion CCR_{OO} and BCC_{OO} within all approaches, are thoroughly discussed below, at first linear CCR_{OO}:

$$Min \sum_{i=1}^{m} v_i x_{ip} + \sum_{d=1}^{D} w_d z_{dp} \tag{4}$$

$$s.t. \sum_{r=1}^{s} u_r y_{rp} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \sum_{d=1}^{D} w_d z_{dj} \leq 0 \quad , j = 1, \dots, n$$

$$u_r, v_i, w_d \geq \varepsilon$$

Dual proposed model in CCR$_{OO}$:

$$Max \; \varphi_p$$

$$s.t. \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ip} \qquad , i = 1, \dots, m$$

$$\sum_{j \overline{n} 1}^{n} \lambda_j z_{dj} \leq z_{dp} \qquad , d = 1, \dots, D$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi_p y_{rp} \qquad , r = 1, \dots, s \tag{5}$$

$$\lambda_j \geq 0, \varphi_p \; free \quad j=1,\dots,n$$

Linear proposed model in BCC$_{OO}$:

$$Min \sum_{i=1}^{m} v_i x_{ip} + \sum_{d=1}^{D} w_d z_{dp} + w = 1$$

$$s.t. \sum_{r=1}^{s} u_r y_{rp} = 1 \tag{6}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \sum_{d=1}^{D} w_d z_{dj} + w \leq 0 \quad , j = 1, \dots, n$$

$$u_r, v_i, w_d \geq \varepsilon, w \; free$$

Dual proposed model in BCC$_{OO}$:

$$Max \; \varphi_p$$

$$s.t. \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ip} \qquad , i = 1, \dots, m$$

$$\sum_{j=1}^{n} \lambda_j z_{dj} \leq z_{dp} \qquad , d = 1, \dots, D$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi_p y_{rp} \qquad , r = 1, \dots, s \tag{7}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0, \varphi_p \; free \quad j=1,\dots,n$$

4.3. Evaluation in window analysis

When there is a restricted number of DMUs, DEA window analysis enables the comparison of a DMUs efficiency in a particular period. Also, it is relevant when there is an emphasis on analyzing the changes in efficiency over time. Figure 2 elaborates the proposed model.
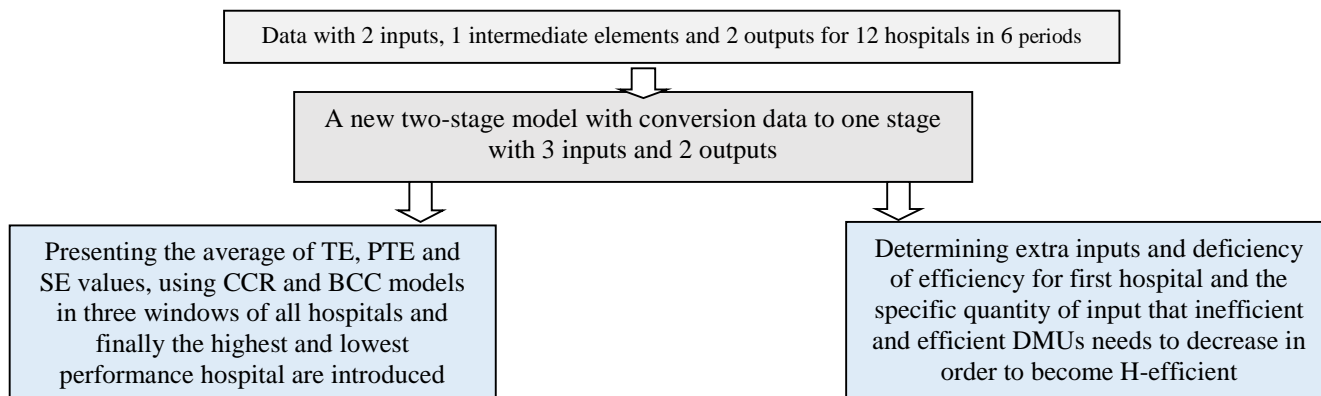

Fig. 2 Implementation of DEA for efficiency analysis

5. Result and discussion

The window analysis reflects each driving hospital as a diverse hospital in each of the splits listed at the top of the table so as to get the scores itemized in the rows that establish the window, whereas the remains on the left side specify the window span and the periods enclosed. For instance, the first row extends from 2011 to 2014 for a window length of four halves that is displayed in the first row. The next row starts in the second year (2012) and prolongs to 2015 which characterizes another window and so on. This results in three windows. Then, there are different data points to which the DEA model is applied to obtain the efficiency scores. Firstly, the Technical Efficiency (TE) and Pure Technical Efficiency (PTE) are computed using CCR and BCC models, respectively. The Scale Efficiency (SE) is then calculated. Tables II present the TE, PTE and SE values for the first hospital, respectively. To illustrate, the first hospital in the first window extends from 2011 to 2014 for a window length of 4 years, in 2014, TE (= 0.752), PTE (= 0.756) and finally, SE is calculated as follows:

$$SE=TE/PTE=0.752/0.756=0.994 \tag{8}$$

TABLE II. Window analysis for TE in the first hospital

TE	2011	2012	2013	2014	2015	2016	Avg.	Std.	CV	Min	Max
2011-2014	1	1	1	0.752			0.938	0.124	0.132	0.752	1
2012-2015		1	1	0.789	1		0.947	0.105	0.11	0.789	1
2013-2016			1	0.924	0.956	1	0.97	0.037	0.038	0.924	1
Avg.	1	1	1	0.821	0.978	1					

TABLE III. Window analysis for PTE in the first hospital

PTE	2011	2012	2013	2014	2015	2016	Avg.	Std.	CV	Min	Max
2011-2014	1	1	1	0.756			0.939	0.122	0.129	0.756	1
2012-2015		1	1	0.837	1		0.959	0.081	0.084	0.837	1
2013-2016			1	1	1	1	1	0	0	1	1
Avg.	1	1	1	0.864	1	1					

TABLE IV. Window analysis for SE in the first hospital

SE	2011	2012	2013	2014	2015	2016	Avg.	Std.	CV	Min	Max
2011-2014	1	1	1	0.994			0.998	0.003	0.003	0.994	1
2012-2015		1	1	0.942	1		0.985	0.029	0.029	0.942	1
2013-2016			1	0.924	0.956	1	0.97	0.037	0.037	0.924	1
Avg.	1	1	1	0.953	0.978	1					

Applying the aforesaid arrangement of H-efficient (100%), Efficient (between 90% and 100%) and Inefficient (less than 90%) means that all three windows in TE are Efficient. The first two windows of PTE are Efficient and the last window of PTE is H-efficient. All three windows in SE are Efficient. The columns and rows of window analysis deliver two discrete purposes:

- The columns in window DEA are used to inspect stability properties. The TE, PTE and SE values of the first hospital in Table II showed a stable performance for two periods. In fact, TE, PTE, and SE in 2012 and 2013 are set to 1 for three windows.
- The row window is scrutinized to ascertain the presentation trends. For example in the first row, the TE average is 0.938 and a standard deviation is set to 0.124. It is found that the Coefficient of Variation (CV) for TE values in all windows except the last window for the first hospital is larger than 5%, which indicates the presence of trends in the efficiencies of the same window.

The results in Table V, VI, VII, and VIII are used for determining the scarcity of outputs and deficiency of efficiency for the first hospital and the specific quantity of output that inefficient and efficient DMUs needs to increase in order to become H-efficient. In fact, actual, target and potential improvement of DMUs which are not H-efficient in each window, are introduced in order to stand on the efficiency frontier line for the smallest and largest TE and PTE average values. Unless largest PTE average, which is H-efficient in that actual is set to target and potential improvement, other averages need to increase the number of outputs.

Table V. The only inefficient DMU (0.752) in 2014 in the smallest TE Avg. (0.938)

Variable	Actual	Target	Improvement
Output1(OT)	16100	21422	+33.06%
Output2(IT)	1800	4707	+161.49%

Table VI. The first efficient DMU (0.924) in 2014 in the largest TE Avg. (0.97)

Variable	Actual	Target	Improvement
Output1(OT)	16100	17426	+8.24%
Output2(IT)	1800	3615	+100.85%

Table VII. The second efficient DMU (0.956) in 2015 in the largest TE Avg. (0.97)

Variable	Actual	Target	Improvement
Output1(OT)	17693	18507	+4.6%
Output2(IT)	4500	4708	+4.6%

Table VIII. The only inefficient DMU (0.756) in 2014 in the smallest PTE Avg. (0.939)

Variable	Actual	Target	Improvement
Output1(OT)	16100	21306	+32.33%
Output2(IT)	1800	5019	+178.85%

From Tables II, III, IV, V, VI, VII, and VIII the following results are obtained:
The first window of TE for the first hospital (2011-2014) created the smallest TE average value (= 0.938). The TE for the only DMU which is not H-efficient in 2014 (= 0.752) suggests that the same level of output could be produced with 75.2% of the resources if this DMU was acting out on the Frontier. It could be inferred also that 24.08% of overall resources could be saved by rising the performance of this DMU to the highest level. In Table V, in order for the first hospital in this DMU to become H-efficient, it needs to rise 5322 or 33.06% units of the OT and increase 2907 or 161.49% of IT. On the one hand, other DMUs of this window stands on the efficiency frontier line and does not need to increase any of the outputs. All above-mentioned steps can be inferred from the largest TE and the smallest and the largest PTE. Table IX displays an arrangement of TE, PTE, and SE efficiency values in each window for all hospitals.

TABLE IX. Average efficiency of each window for TE, PTE, and SE in all hospitals

1ST hospital	TE	PTE	SE	2nd hospital	TE	PTE	SE	3rd hospital	TE	PTE	SE
2011-2014	0.938	0.939	0.998	2011-2014	0.864	0.927	0.926	2011-2014	0.949	0.955	0.992
2012-2015	0.947	0.959	0.985	2012-2015	0.866	0.907	0.950	2012-2015	0.918	0.955	0.959
2013-2016	0.970	1	0.970	2013-2016	0.926	0.952	0.971	2013-2016	0.958	1	0.958
4th hospital	TE	PTE	SE	5th hospital	TE	PTE	SE	6rd hospital	TE	PTE	SE
2011-2014	0.768	0.939	0.805	2011-2014	0.639	0.724	0.811	2011-2014	0.966	0.990	0.974
2012-2015	0.856	0.969	0.951	2012-2015	0.717	0.863	0.868	2012-2015	0.931	0.967	0.959
2013-2016	0.834	1	0.834	2013-2016	0.632	0.792	0.792	2013-2016	0.973	1	0.973
7th hospital	TE	PTE	SE	8th hospital	TE	PTE	SE	9th hospital	TE	PTE	SE
2011-2014	1	1	1	2011-2014	0.999	1	0.999	2011-2014	1	1	1
2012-2015	1	1	1	2012-2015	1	1	1	2012-2015	1	1	1
2013-2016	1	1	1	2013-2016	0.916	1	0.916	2013-2016	0.919	0.921	0.922
10th hospital	TE	PTE	SE	11th hospital	TE	PTE	SE	12th hospital	TE	PTE	SE
2011-2014	1	1	1	2011-2014	0.997	1	0.997	2011-2014	0.988	1	0.988
2012-2015	0.926	0.931	0.933	2012-2015	1	1	1	2012-2015	0.927	1	0.927
2013-2016	0.891	0.913	0.966	2013-2016	1	1	1	2013-2016	1	1	1

Disintegrating technical efficiency scores into PTE and SE offers regulation on what can be attained in the short and long terms. If the widely held of inefficiency in any hospital is due to low value of SE which means a small size of medical operation, then the hospital needs to be expanded. Alternatively, the PTE value can be frequently addressed in the short term without changing the scale of medical operations. The low value of PTE is due to managerial deficit problem in organizing the inputs in the medical process. It should be noted that it is easier to increase the TE than to increase SE. When a hospital becomes technically efficient, it makes sense to deal with SE, so hospital manager should focus on growing the TE. For example, in Table IX, the average efficiency of TE for the fourth hospital in the first window (2011-2014) with 0.768 is considered inefficient because of low SE (= 0.805). PTE in this window is 0.939. Then, the low TE value in this window is attributed to the low performance in scale efficiency and the expansion is the solution in this case. Moreover, the average TE for the fifth hospital in the first window of 0.639 is inefficient because of low PTE (= 0.724). SE in this window is 0.811. Then, the low TE value in this window is attributed to the low performance in PTE due to the low operation of the inputs or a problem in converting inputs into outputs and the number the outpatient treated (OT) and inpatient treated (IT) in the hospitals should be increased. The efficiency groupings and average efficiency of each window for TE, PTE, and SE for all hospitals, corresponds to each category are considered and finally, Seventh hospital with H-efficient category in all windows of TE, PTE and SE is considered as the highest performance hospital and Fifth hospital with inefficient category in all windows of TE, PTE and SE is considered as the lowest.

6. Conclusion

Various heart hospitals efficiency in DEA were compared. A unique model consisting of an exclusive model and converting two-stage model to a novel one stage model and evaluation methods are developed. Output-oriented window analysis with TE, PTE, and SE are applied to find the superior hospitals and periods. Hospitals are compared in terms of their actual, target and potential improvement. The highest and lowest performance hospitals within various approaches are introduced. The research can help by means of a practical tool for the general practitioner to efficiently associate efficiency of indeterminate cases and advise consequently.

7. References

[1] S. Mendis, P. Puska, B. Norring, "World Health Organization Global Atlas on Cardiovascular Disease Prevention and Control," 2011.

[2] M.S.C Lobo, Y Ozcan, M.P.E Lins, A.C.M Silva, R. Fiszman, "Financing reform and productivity change in Brazilian teaching hospitals: Malmquist approach," Central European Journal of Operations Research 18:141–152, 2010.

[3] M. Caballer-Tarazona, I. Moya-Clemente, D. Vivas-Consuelo and I. Barrachina-Martínez, " A model to measure the efficiency of hospital performance," *Mathematical and Computer Modelling* 52: 1095-1102, 2010.

[4] M. Mirmozaffari, A. Alinezhad, and A. Gilanpour, "Data Mining Classification Algorithms for Heart Disease Prediction," Int'l Journal of Computing, Communications & Instrumentation Engg (IJCCIE), ISSN 2349-1469 EISSN 2349-1477, Vol.4, Issue1, Jan 2017.

[5] M. Mirmozaffari, A. Alinezhad, and A. Gilanpour, "Heart Disease Prediction with Data Mining Clustering Algorithms," Int'l Journal of Computing, Communications & Instrumentation Engg (IJCCIE), ISSN 2349-1469 EISSN 2349-1477, Vol.4, Issue1, Jan 2017.

[6] M. Mirmozaffari, A. Alinezhad, and A. Gilanpour, "Data Mining Apriori Algorithm for Heart Disease Prediction," Int'l Journal of Computing, Communications & Instrumentation Engg (IJCCIE), ISSN 2349-1469 EISSN 2349-1477, Vol.4, Issue1, Jan 2017.

[7] M. Mirmozaffari, A. Alinezhad. "Ranking of Heart Hospitals Using Cross-Efficiency and Two-stage DEA," 7th International Conference on Computer and Knowledge Engineering (ICCKE 2017), October 26-27 2017, Ferdowsi University of Mashhad, 978-1-5386-0804-3/17/$31.00 ©2017 IEEE.

[8] C. Kao and S.N. Hwang, (2008), Efficiency Decomposition in Two-Stage Data Envelopment Analysis: an Application to non-Life Insurance Companies in Taiwan. European Journal of Operational Research 185, 418-429.

[9] Y.M. Wang, and K.S. Chin, (2010) Some Alternative DEA Models for Two-stage Process. Expert Systems with Applications 37, 8799-8808.

[10] F. Hosseinzadeh Lotfi, A. Toloie Eshlaghy, M. shafiee, H. Saleh1, H. Nikoomaram and S. M. Seyedhoseini, A new two-stage data envelopment analysis (DEA) model for evaluating the branch performance of banks, African Journal of Business Management Vol. 6(24), pp. 7230-7241, 20 June 2012.