



Non-discretionary Factors in Data Envelopment Analysis: Review and Extension

Mehdi Fallah Jelodar*

^(a) *Department of Mathematics, Ayatollah Amoli branch, Islamic Azad University, Amol, Iran.*

Received August 24, 2016, Accepted November 29, 2106

Abstract

Data Envelopment Analysis (DEA) technique uses linear programming to evaluate the relative efficiency of a homogeneous set of Decision Making Units (DMUs) in their use of multiple inputs to produce multiple outputs. The standard DEA models do not take into account non-discretionary inputs and outputs and ignore the possibility that efficiency may be correlated with the non-discretionary factors. However, one key issue in performance measurement problems is how to treat non-discretionary factors, which influence the performance of DMUs and are, at the same time, out of the control of the management. In this paper, a new model for measuring efficiency is defined such that non-discretionary factors are taken into account by the decision maker. The main contributions of this paper are fourfold: (1) we review the existing approaches for measuring efficiency scores to control non-discretionary factors in production; (2) we provide a discussion of strengths and weaknesses and highlighting potential limitations of the existing non-discretionary DEA models; (3) we propose a new approach based on relative importance of non-discretionary inputs that overcomes existing weaknesses; (4) we use a numerical example to demonstrate the feasibility and richness of the obtained solutions.

Keywords: Data Envelopment Analysis, Efficiency, Non-discretionary Factors.

*. Corresponding author, Email address: m.fallahjelodar@iauamol.ac.ir

1. Introduction

Data Envelopment Analysis (DEA), initially introduced by Charnes et al. [5] and extended by Banker et al. [2], is a well known non-parametric methodology for computing the relative efficiency of a set of homogeneous units, named Decision Making Units (DMU). The non-parametric property implies that this methodology does not rely on assumptions requiring the data to follow from any specific production function. DEA uses the data observed and some preliminary assumptions to determine a production possibility set which contains those operating points that are deemed feasible. Then, DEA formulates and solves a linear programming (LP) problem that produces an efficiency score and a target operating point for each DMU. The target operating point lies on the efficient frontier and is computed in such a way that it generally uses the same or less inputs to produce identical or more output. The efficiency score is a measure of the relative improvements in inputs and outputs that can be defined between the DMU and its assigned target.

Traditional DEA models assume that the measured units are homogeneous. In other words, they carry out the same ability with similar objectives, use similar inputs and produce similar outputs, and run in similar operational environments. Sometimes the supposition of homogeneous environments is disregarded and factors that describe the differences in the environments need to be included in the analysis. These factors, and others outside the control of the DMUs, are frequently called non-discretionary factors. The standard DEA models do not take into account non-discretionary inputs and outputs and ignore the possibility that efficiency may be correlated with the non-discretionary factors. Thus, one key issue in performance measurement problems is how to treat non-discretionary factors, which influence the performance of DMUs and are, at the same time, out of the control of the management. Usually, the management

can decide on some controllable factors internal to production activities, while the impact of the operating environment is out of the control of the management. Therefore, several researches developed different models for considering non-discretionary factors in DEA models. These approaches are developed for controlling the non-discretionary inputs. The DEA model is coined by Banker and Morey [3] for fulfilling what above said. Convexity is an assumption by considering either discretionary or non-discretionary inputs. These classes of inputs were treated differently, however, by not allowing radial reduction in the non-discretionary inputs. Ruggiero [17] extended this model by dropping the convexity constraint associated with the non-discretionary inputs. Rather, non-discretionary inputs were treated as shift exclude DMUs with more favorable levels of the non-discretionary factor. The approach that is considered here as a third one is introduced by Ray [16], which does not consider the non-discretionary inputs in the DEA model in the first stage. The non-discretionary inputs are controlled in the second stage of regression, which permits an adjusted measure of technical efficiency to enter the model. A hybrid model is announced by Ruggiero [18] that have three stages for allowing the multiple non-discretionary inputs to be paid attention. Simulation analysis [18] showed that multiple stage models of Ray (1991) [16] and Ruggiero (2004) [18] have a superior level in comparison with Banker and Morey model, and are acted better.

In order to compare Banker and Morey model with the stochastic frontier model with one non-discretionary variable, a simulation analysis is used by Yu [19]. The cross-sectional stochastic frontier approach has been depicted by Ondrich and Ruggiero [15] to be of limited value since it does not really allow measurement error. Other concluded result by Yu are consistent with Ruggiero [17]. Moreover, a revised model

has been proposed that produce an undistorted efficiency measure by Ruggiero [18]. As discussed in that paper and illustrated with simulation analysis, the performance of the existing model declines as the relationship between non-discretionary inputs and true but unobserved efficiency gets stronger. In addition to discussing the problem, that paper introduced a new DEA model which overcomes the identified problems. One shortcoming, however, was the reliance on parametric techniques to identify this relationship. Hosseinzadeh Lotfi et al. [11] discussed and reviewed the use of super-efficiency approach in data envelopment analysis (DEA) sensitivity analyses the presence of non-discretionary inputs. Camanho et al. [4] proposed an enhanced DEA model that accommodates non-discretionary inputs and outputs and treats them differently depending on their classification as internal or external to the production process. Ebadi and Shiri Shahraki [6] extended the definition of return to scale and scale elasticity when some inputs and outputs were non-discretionary. Moreover, they presented an efficient algorithm, based upon a simplex algorithm, to determine scale elasticity in the existence of non-discretionary factors. Jahanshahloo et. al. [12] used a non linear form of a non-radial DEA model to consider non-discretionary factors. Gholam Abri and Fallah Jelodar [10] extended their methods to break the existing weaknesses and proposed a linear model. Azizi and Ganjeh Ajirlu [1] proposed a novel pair of DEA models for measurement of relative efficiencies of DMUs (DMUs) in the presence of non-discretionary factors and imprecise data. Khoshandam et al. [13] introduced a DEA approach to calculate marginal rates of substitutions between discretionary inputs/outputs and non-discretionary outputs. In this paper, a new model for measuring efficiency is defined such that non-discretionary factors are taken

into account by the decision maker. The main contributions of this paper are fourfold: (1) we review the existing approaches for measuring efficiency scores to control non-discretionary factors in production; (2) we provide a discussion of strengths and weaknesses and highlighting potential limitations of the existing non-discretionary DEA models; (3) we propose a new approach based on relative importance of non-discretionary inputs that overcomes existing weaknesses; (4) we use a numerical example to demonstrate the feasibility and richness of the obtained solutions.

The rest of this paper is organized as follows. Section 2 first introduces the basic DEA model for measurement of efficiencies of DMUs and then reviews the existing models for controlling non-discretionary inputs together with providing their potential strengths and weaknesses. Based on this discussion, a new method is developed that handles multiple non-discretionary factors. Section 3 presents a numerical example to illustrate the application of the proposed model. Conclusions are set forth in Section 4.

2. DEA background

2.1. DEA models without non-discretionary inputs

The aim of this section is to review DEA models without non-discretionary inputs for evaluating the efficiencies of DMUs.

Consider n DMUs with m inputs and s outputs. The input and output vectors of DMU_j ($j=1, \dots, n$) are $x_j = (x_{1j}, \dots, x_{mj})^t$ and $y_j = (y_{1j}, \dots, y_{sj})^t$, respectively, where $x_j \geq 0, x_j \neq 0, y_j \geq 0, y_j \neq 0$. By using the variable return to scale, convexity and possibility postulates, the non-empty production possibility set (PPS) is defined as follows:

$$T_v = \left\{ (x, y) : x \geq \sum_{j=1}^n \lambda_j x_j, y \leq \sum_{j=1}^n \lambda_j y_j, \right. \\ \left. \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n \right\}$$

Based on this definition, the BCC model proposed by Banker et al. (1984) [2] and based on the work of Farrell (1957) [9] is as follows:

$$\begin{aligned} F(x_o, y_o) = \min \quad & \theta \\ \text{s.t.} \quad & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \quad i = 1, \dots, m \\ & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \dots, s \\ & \sum_{j=1}^n \lambda_j = 1 \\ & \lambda_j \geq 0, \quad j = 1, \dots, n \end{aligned} \quad (1)$$

The dual problem of Model (1) is given as follows:

$$\begin{aligned} DF(x_o, y_o) = \max \quad & \sum_{r=1}^s u_r y_{ro} - u_0 \\ \text{s.t.} \quad & \sum_{i=1}^m v_i x_{io} = 1 \\ & \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} - u_0 \leq 0, \quad j = 1, \dots, n \\ & u_r \geq 0, \quad r = 1, \dots, s \\ & v_i \geq 0, \quad i = 1, \dots, m \end{aligned} \quad (2)$$

Following the work of Lovell (1993) [14], the above-mentioned production technology transform inputs $x = (x_1, \dots, x_m)$ into output $y = (y_1, \dots, y_s)$ for $j = 1, \dots, n$ firms can be represented with the input set:

$$L(y) = \{x : (x, y) \text{ is feasible}\}.$$

For every output vector y , $L(y)$ has isoquant as

$$IsoqL(y) = \{x : x \in L(y), \lambda x \notin L(y), \lambda \leq 1\}$$

and efficient subset as

$$Eff(y) = \{x : x \in L(y), x' \notin L(y), x' \leq x\}.$$

It is important to note that the radial Farrell measure does not require comparison of a

given input vector to an input vector that belong to the estimated efficient subset.

2.2. Non-discretionary DEA models and a new model

The traditional DEA models assume that all inputs are discretionary. Now assume that each DMU uses a vector x of inputs to produce a vector y of outputs given x vector non-discretionary inputs $z = (z_1, \dots, z_k)$.

These non-discretionary inputs affect on the transformation of discretionary inputs into outputs. For convenience, the vector z is defined so that increases in any component leads to a more favorable environment, ceteris paribus. The production technology transforms input vector x into output vector y can be represented by the conditional input set:

$$L(y | z) = \{x : (x, y) \text{ is feasible for given } z\}$$

For every output vector y , $L(y)$ has isoquant as

$$IsoqL(y | z) = \{x : x \in L(y | z), \lambda x \notin L(y | z), \lambda \leq 1\}$$

$$and \text{ efficient subset } Eff(y | z) = \{x : x \in L(y | z), x' \notin L(y | z), x' \leq x\}.$$

Note that $L(y | z) \subseteq L(y | z')$ implies that z' is more favorable environment than z . Given multiple non-discretionary inputs, it is necessary to identify the importance of each non-discretionary factor in production process.

The first DEA model to allow continuous non-discretionary variables was developed by Banker and Morey [3]. Recognizing the inappropriateness of treating fixed factors a discretionary, the authors modified the constraints on the fixed inputs. The BM input oriented (variable return to scale) efficiency measure for production possibility (x_o, y_o) is as follows:

$$BM(x_o, y_o) = \min \quad \theta$$

$$\text{s.t.} \quad \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \quad i = 1, \dots, m$$

$$\sum_{j=1}^n \lambda_j z_{ij} \leq z_{io}, \quad i = 1, \dots, k$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \dots, s \quad (3)$$

$$\sum_{j=1}^n \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, n$$

The dual problem of Model (3) is given as follows:

$$DBM(x_o, y_o) =$$

$$\max \sum_{r=1}^s u_r y_{ro} - \sum_{i=1}^k w_i z_{io} - u_0$$

$$S.t. \sum_{i=1}^m v_i x_{io} = 1$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} - \sum_{i=1}^k w_i z_{ij} - u_0 \leq 0, \quad (4)$$

$$j = 1, \dots, n$$

$$u_r \geq 0, \quad r = 1, \dots, s$$

$$v_i \geq 0, \quad i = 1, \dots, m$$

$$w_i \geq 0, \quad i = 1, \dots, k$$

The constraints on fixed factors are similar to the constraints on discretionary inputs; they are modified, however, to break the link between efficiency and fixed factors. This modification purportedly controls for fixed factors of production by requiring a convex combination of the referent production possibilities to have an environment no better than the DMU under analysis. Ruggiero [17], however, showed that the referent production possibility may not be feasible, because return to scale should be defined relatively only to discretionary inputs. Enforcing convexity with respect to the non-discretionary inputs leads to improper restriction of the production possibility sets and distorted efficiency measurement.

To evaluate a given DMU, it is necessary to exclude DMUs with a more favorable environment. This was achieved with the public sector model of Ruggiero [18]. The Ruggiero input oriented (variable return to scale) efficiency measure for production possibility (x_o, y_o) is as follows:

$$R1(x_o, y_o) = \min \theta$$

$$S.t. \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \quad i = 1, \dots, m$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \dots, s$$

$$\lambda_j = 0, \quad \text{if } \exists z_{ij} > z_{io}, \quad i = 1, \dots, k \quad (5)$$

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

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

This model explicitly restricts the comparison set to exclude DMUs that face a more favorable environment. Similar to the BM model, this model requires a priori specification of the continuous non-discretionary variables. Importantly, as the numbers of continuous fixed factors increases, the probability of identifying a DMU as efficient by default increases. This ignores comparisons between a given DMU and another DMU that overall, has the same or worse environment even though it has a more favorable level of at least one non-discretionary input. This fact suggests an inherent weakness of the Ruggiero model. To remove these weaknesses, Ruggiero [17] modified his model and proposed the following linear programming to measure technical efficiency in the presence of non-discretionary factors:

$$R2(x_o, y_o) = \min \theta$$

$$S.t. \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \quad i = 1, \dots, m$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \dots, s \quad (6)$$

$$\lambda_j = 0, \quad \text{if } Z_j > Z_o, \quad \forall j = 1, \dots, n$$

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

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

Model (6) prevents DMUs with a higher level of the non-discretionary input into

reference set. One key assumption in model (6) is that true efficiency is not correlated with non-discretionary factors. As shown in Ruggiero [17], model (6) has some weaknesses. The problem arises because non-discretionary factors has two effects on production: it simultaneously determines the location of the true frontier and effects the distance from the frontier. The efficiency measure $R2(x_o, y_o)$ of model (6) is unable to disentangle the two effects, attributing both effects to the location of the frontier.

To remove the difficulties of pervious models, we propose a two stage model. In first stage use model (4) for evaluating all DMUs. Then $\sum_{i=1}^k w_i^* z_{ij}$ is the "relative importance" of non-discretionary factors obtained by DBM model. In this case, the following linear programming is proposed to obtain true efficiency of DMUs:

$$\theta^{ND} = \min \theta$$

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

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, \dots, s$$

$$\sum_{j=1}^n \lambda_j = 1, \quad \lambda_j \geq 0,$$

$$\text{if } \sum_{i=1}^k w_i^* z_{ij} < \sum_{i=1}^k w_i^* z_{io}, \quad (7)$$

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

Problem may be occurred when problem (7) has alternative optimal solutions. In this case one may use especial measures to choose one of the solutions. For example, one may use lexico minima of the vector of optimal solutions.

3. A numerical example

In this section, we present a numerical example to illustrate the applicability and efficiency of the proposed model. Consider 20 iranian bank branches with two non-discretionary inputs and three discretionary inputs and four outputs. The first non-discretionary input is the area of branch and the second is score of staff's education. Normalized data is used to illustration. We added Ruggiero's model (model (6)) and proposed model (model (7)) through these data. These data and results are summarized in Table 1:

Table 1: Data and results of the example

DMUs	N-D inputs		D inputs			Outputs				Ruggiero	$\sum_{i=1}^2 w_i^* z_{ij}$	New model
	Z_1	Z_2	I_1	I_2	I_3	O_1	O_2	O_3	O_4			
1	0.9333	0.7006	0.8276	0.4404	0.1820	0.0556	0.1077	0.0026	0.5803	0.5804	16E-6	0.7105
2	0.8667	0.9641	0.6810	0.3793	0.1541	0.0518	0.0684	0.0107	0.6871	0.9166	18E-6	1
3	1.0000	0.8862	0.7974	0.3301	0.1030	0.0857	0.1394	0.0039	0.7942	0.8357	19E-6	0.9341
4	1.0000	0.4371	1.0000	0.2058	0.3468	0.2213	0.1611	0.0064	1.0000	0.8635	0.680576	0.9543
5	0.9333	0.7964	0.7888	0.3601	0.1966	0.0949	0.0817	0.0102	0.7922	0.89166	17E-6	0/9687
6	0.9333	1.0000	0.7759	0.4424	0.0000	0.0644	0.0330	0.0038	0.5137	0.7852	19E-6	0.8496
7	0.7333	0.5329	0.6250	0.2579	0.0000	0.0454	0.0890	0.0023	0.5449	1	13E-6	1
8	0.6667	0.7365	0.5733	0.1102	0.0005	0.0253	0.0514	0.0050	0.3229	1	14E-6	1
9	0.6667	0.8563	0.5086	0.2552	0.3636	0.0698	0.1010	0.0007	0.6304	1	15E-6	1
10	0.5333	0.3413	0.5086	0.1698	0.0000	0.0778	0.1401	0.0172	0.4081	1	9E-7	1
11	0.8000	0.6467	0.6940	0.2675	0.0355	0.0687	0.0544	0.0011	0.8005	0.9345	14E-6	1
12	0.6000	0.6228	0.5043	0.0514	0.0011	0.0710	0.0743	0.0031	0.4810	1	12E-6	1
13	0.6000	0.6168	0.5129	0.0617	0.0206	0.0724	0.0447	0.0010	0.3620	0.7221	12E-6	0.7221
14	0.7333	0.6347	0.6207	0.1260	0.1383	0.0591	0.0723	0.0011	0.4404	0.6165	14E-6	0.6971
15	0.6667	0.7725	0.5431	0.0674	0.0044	0.0537	0.0663	0.0124	0.5679	1	14E-6	1
16	0.6000	0.4850	0.5302	0.1461	0.1000	0.0395	0.0532	0.0068	0.4320	0.7811	13E-6	0.8412
17	0.8000	0.4731	0.7672	0.1828	0.0028	0.4035	0.1294	0.0026	0.9467	0.9321	13E-6	1
18	0.4000	0.2575	0.3707	0.0823	0.0043	0.0323	0.0174	0.0007	0.4249	1	7E-7	1
19	0.9333	0.6287	0.8405	0.3395	0.5039	0.1402	0.3263	0.0207	0.6414	1	16E-6	1
20	0.6667	0.3892	0.6336	0.1749	0.0023	0.0911	0.2292	0.0015	0.7404	1	11E-6	1

Following to this example all non-discretionary factors should be consider in evaluating efficiency score of each DMU by using Ruggiero's model. But by using the proposed model the efficiency score of DMUs are improved, therefore, this example show that some of the non-discretionary factors have not any correlation with efficiency score but they considered in Ruggiero's model. Also the tradition model did not obtain true efficiency score. For example by considering DMU 4, it is clear that all non-discretionary inputs are considered for evaluating its efficiency score in Ruggiero's model, but by using the new model the non-discretionary inputs of other DMUs dose not have any effect on efficiency score of this unit, therefore, they did not consider in evaluation and its efficiency score is improved from 0.8635 to 1.

4. Conclusion

The traditional DEA models do not take into account non-discretionary inputs and outputs and ignore the possibility that

efficiency may be correlated with the non-discretionary factors. Thus, how to treat non-discretionary factors, which influence the performance of DMUs and are, at the same time, out of the control of the management is an important issue in performance measurement problems. In this paper we reviewed some basic non-discretionary DEA models together their strengths and weaknesses. Then, we proposed a new model based on relative importance of non-discretionary inputs that overcomes existing weaknesses. Finally, we used a numerical example to demonstrate the feasibility and richness of the obtained solutions. In future research, the proposed method should be developed to consider strategic environments in which the values of inputs and outputs are uncertain. Moreover, further research should be performed to develop the proposed non-discretionary DEA model in the present of undesirable variables.

References

- [1] Azizi, H., Ganjeh Ajirlu, H., (2011). Measurement of the worst practice of decision-making units in the presence of non-discretionary factors and imprecise data, *Applied Mathematical Modelling* 35, 4149–4156.
- [2] Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30 (9), 1078-1092.
- [3] Banker, R.D, Morey, R.C, 1986, Efficiency analysis for exogenously fixed inputs and outputs. *European Journal of Operational Research* 553-565.
- [4] Camanho, A. S., Portela, M.C., Vaz, C.B., (2009). Efficiency analysis accounting for internal and external non-discretionary factors. *Computers & Operations Research* 36, 1591 – 1601.
- [5] Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 429-444.
- [6] Ebadi, M.J., Shiri Shahraki, M.S., (2010). Determination of scale elasticity in the existence of non-discretionary factors in performance analysis, *Knowledge-Based Systems*, 23, 434–439.
- [7] Fare, R., Grosskopf, S., Lovell, C.A.K., 1994. *Production frontiers*, Cambridge University Press, New York.
- [8] Fare, R., Grosskopf, S., Lovell, C.A.K., 1985. *The management of efficiency of production*. Kluwer-Nijhoff, Boston, MA.
- [9] Farrell, M.J, 1957. The measurement of productive efficiency. *Journal of the royal statistical science series A*, General 120, 253-281.
- [10] Gholam Abri, A., Fallah Jelodar, M., 2012. A complementary on proposing a new model on data envelopment analysis by considering non-discretionary factors and a review on previous models. *Mathematical and Computational Applications*, 17(3), 244-246.
- [11] Hosseinzadeh Lotfi, F., Jahanshahloo, G.R., Esmaili, M., 2007. Sensitivity analysis of efficient units in the presence of non-discretionary inputs. *Applied Mathematics and Computation*, 190, 1185–1197.
- [12] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Shoja, N., Gholam Abri, A., Fallah Jelodar, M., Jamali firouzabadi, k. 2010. Proposing a new model on data envelopment analysis by considering non-discretionary factors and a review on previous models. *Mathematical and Computational Applications*, 15(3), 344-353.
- [13] Khoshandam, L. Amirteimoori A., Kazemi Matin, R., (2014). Marginal rates of substitution in the presence of non-discretionary factors: A data envelopment analysis approach, *Measurement* 58, 409–415.
- [14] Lovell, C.A.K, 1993. *Production frontiers and productive efficiency*. Oxford university press, New York, pages 3-67.
- [15] Ondrich, J., Ruggiero, J., (2001). Efficiency measurement in the stochastic frontier model. *European Journal of Operational Research*, 129(2), 434-442
- [16] Ray, S.c., 1991. Resource use efficiency in public schools. A study of connecticut data, *Management Science* 37, 1620-1628.
- [17] Ruggiero, J., 1996. On the measurement of technical efficiency in the public sector, *European Journal of Operational Research* 553-565.
- [18] Ruggiero, J., 2004. Performance evaluation when non-discretionary factors correlate with technical efficiency. *European Journal of Operational Research*, 159, 250-257.
- [19] Yu, C, 1998. The effects of exogenous variables in efficiency measurement-A Mont Carlo study. *European Journal of Operational Research* 105, 569-580.