Hacettepe Journal of Mathematics and Statistics
Volume 41 (2) (2012), 321 – 335

A NEW RESTRICTED MODEL USING CORRELATION COEFFICIENTS AS AN ALTERNATIVE TO CROSS-EFFICIENCY EVALUATION IN DATA ENVELOPMENT ANALYSIS

Emine Demet Mecit*† and İhsan Alp‡

Received 11 : 04 : 2011 : Accepted 10 : 10 : 2011

Abstract

In this paper, a new Data Envelopment Analysis (DEA) model, which is restricted using correlation coefficients (the CCRCOR), was compared with the cross-efficiency evaluation model (the CEM) and some other models. We used two hypothetical data sets, which were generated by E. Thanassoulis (*A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments*, J. Operational Research Society **44**, 1129–1144, 1993) (with one input) and K. F. Lam (*In the determination of weight sets to compute cross-efficiency ratios in DEA*, J. Operational Research Society **61**, 134–143, 2010) (with two inputs). Until now, these data sets have been used by various researchers. The writers did not consider whether or not these data sets were appropriate to the analysis. So, we recommend the CCRCOR model in our paper. The CCRCOR model allows one to add weight restrictions to the CCR model, which are obtained using correlations between input and output variables.

Keywords: Correlation, CCR model, Cross efficiency method, Super efficiency model, True efficiency.

2000 AMS Classification: 62 H 20, 90 C 08.

---

*Hacettepe University, Faculty of Science, Department of Statistics, 06800 Beytepe, Ankara, Turkey. E-mail: `demetm@hacettepe.edu.tr`

†Corresponding Author.

‡Gazi University, Faculty of Sciences, Department of Statistics, Teknikokullar-Beşevler, Ankara, Turkey. E-mail: `ihsanalp@gazi.edu.tr`

1. Introduction

Data Envelopment Analysis (DEA) was developed by Charnes *et al.* [4]. It calculates the efficiency rates as follows:

Efficiency = weighted sum of output / weighted sum of input.

Due to the nature of the classical DEA models, they allow the flexibility of weight for decision-making units (DMUs) under evaluation. Thus, the DMUs work to maximize their own efficiency scores. Thus, the weights of input and output variables are determined. So, some important variables in the optimal solution by the classical DEA can be assigned zero weight values.

However, some unimportant variables in the analysis can take weight values close to one (1). This is one of the disadvantages highlighted in the DEA literature. The determination of the weights without the need for a prior knowledge is an advantage of DEA. Due to the flexibility of weight, the weight factor needs checking against the risk of assigning unrealistic weights in some cases [13]. In the classical DEA, very different weights for the inputs and outputs of the DMU under assessment may be assigned [5]. To avoid large differences in terms of the weight values among all DMUs, the Assurance Region (AR) concept was introduced by Thompson *et al.* [17]. Thus, weight restrictions are often used to destroy the flexibility of weight in DEA. It requires a pre-knowledge of the application by area experts to create these weight restrictions in the AR. But, the cross-efficiency evaluation method (CEM) is a model without the need of prior knowledge. It was first introduced by Sexton *et al.* [14]. Another disadvantage is the problem of ranking in the classical DEA, that is not obtaining a ranking from the efficiency scores of the DMUs. A DMU under evaluation in DEA maximizes its own efficiency score, but it does not take into consideration the other DMUs. So, the DMU's weights are assigned in the classical DEA. But sometimes, it can lead to misleading results. The cross-efficiency score of the assessed DMU is calculated based on the average cross efficiencies obtained by using all weights to overcome these problems. DEA provides self-assessment of each unit, but the CEM provides peer evaluation for the DMU under evaluation. The DMUs in the DEA are divided into two classes as the efficient and inefficient.

Lam [9] comments that one of the deficiencies in DEA may be finding multiple solutions as linear solutions of the efficient DMUs. Thus, the set of weights obtained is only one of many optimal weight sets [9]. When a set of weights is used in cross-evaluation, it comes up as a question why this cluster is preferred rather than others. Another shortcoming of DEA is that it may give some useless or excessively different weight values for the DMUs under evaluation. Also, zero input and output weights are seen frequently in DEA. Lam [9] developed a new method of implementing mixed-integer linear programming. The purpose of the model is to use cross-evaluation to select the appropriate weight sets. One advantage of the proposed method is that each weight set obtained can reflect the relative strengths of the efficient DMU under evaluation [9]. The method tries to protect the results of the original classification and to avoid zero weights.

So far, the idea of taking into account weight restrictions imposed by the correlations between input and output variables has not been into account in the DEA literature. From this point of view, we began this study. Hypothetical data sets (see Thanassoulis [16] and Lam [9]) were used in our study. There were meaningless correlations with low levels between the input and output variables in these data sets. But, this situation was not taken into account by the authors. Thanassoulis [16] used this data set in regression analysis. However, this data set does not meet the basic condition of linear regression analysis. According to this condition, variables must be related with each other. Therefore, the CCRCOR model is superior to the other DEA models in this respect. The CCRCOR model gives weights to the level of relations between input and output variables. Furthermore, the initial correlation matrix is calculated. So, the suitability of this data for DEA will be decided at the beginning of the analysis. Thus, the inclusion of uncorrelated variables in the analysis is prevented. Also, the CCRCOR model eliminates unrealistic solutions caused by the flexibility of the weight in the classical DEA models.

The paper is arranged as follows: Section 1 is the introduction section on the subject of cross efficiency in the literature. We suggest a new model, which is referred to as the CCRCOR model, in Section 2. The models used for comparison in this study are mentioned briefly in the third section. In Section 4, efficiency scores are obtained by applying the methods of the CCRCOR, the CEM, the true efficiency and the super efficiency. Then the methods are compared with other methods, by using the Spearman and the Pearson correlation tests. The results are given in Section 5.

2. A new restricted DEA model (CCRCOR) using correlation coefficients

The input and output variables are related to each other in the production process. In this study, the relationship was reflected in the rate of correlation between the variables. Thus, the weights of the model were found. The weights, which assign the maximum efficiency score of the DMU under evaluation, are chosen randomly in the classical DEA models. In this case, a lot of important inputs and outputs are not considered in determining the efficiency score of the DMU under consideration or, are taken into account for very small amounts. This situation leads to unreasonable results. Many studies in the literature have been done to solve the problem of weighting. Weight balancing studies and the weighted methods may be considered in this context to avoid inconsistent results. For example, in an evaluation study for the efficiencies of universities, weighting research assistants and professors at the same rate of will cause inaccurate results.

The classical DEA evaluates DMUs without taking into account the relative importance of inputs and outputs to each other. In this study, a new model (the CCRCOR model), which adds to the model correlations between input and output variables, is recommended. The weights of the input and output variables, which are very important, should be included at this level in production. The mentioned shortcomings can be overcome by taking into account combinations of the weights. The efficiency scores are calculated at a given level of correlation between the input and output variables in the CCRCOR model. If the current relationship is taken into account in assigning the weights of the variables when calculating the efficiency score with the suggested approach, the weights of the variable are balanced. Other writers working on the balancing of weights give importance only to different weights of 0 and 1. In our opinion, if a balanced concept is based on the degree of importance of a variable in the production process, this variable should be placed with a weight at that level in production. If the weights are created according to this principle, the weights will be "balanced".

The CCRCOR model is an objective model. So, its superiority over subjective methods of the CCRCOR model are as follows: Preference information is not needed. The CCRCOR model is objective, so, the results do not change according to different analysts. The relations between the input and output variables in the CCRCOR model are taken into account from the beginning. So, the CCRCOR model results are quite realistic. Thus, the weights with the CCRCOR model are more homogeneously distributed. Further, the CCRCOR model prevents the determination of the DMU under evaluation as efficient when it is actually inefficient.

The CCRCOR model is formulated as follows:

$$
\begin{aligned}
(2.1)\qquad \max \theta_k = {} & \sum_{r=1}^{s} u_r y_{rk} \\
\text{s.t.}\quad & \sum_{i=1}^{m} v_i x_{ik} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0 \quad (j = 1, \ldots, n) \\
& c_{i,\, i+1} v_{i+1} - v_i \le 0 \quad (i = 1, \ldots, m-1) \\
& p_{i,r}\, u_r - v_i \le 0 \quad (i = 1, \ldots, m), (r = 1, \ldots, s) \\
& b_{r,r+1} u_{r+1} - u_r \le 0 \quad (r = 1, \ldots, s-1) \\
& u_1, u_2, \ldots, u_s \ge 0 \quad , v_1, v_2, \ldots, v_m \ge 0
\end{aligned}
$$

In the CCRCOR model, the meaning of the symbols is as follows:

- θ_k: the efficiency score for the DMU_k,
- u: the output weights,
- v: the input weights,
- x: inputs,
- y: outputs.

Moreover,

- $c_{i,\, i+1}$ is the correlation coefficient between the i^{th} and $(i+1)^{\text{th}}$ input variables.
- $p_{i,\, r}$ is the correlation coefficient between the i^{th} input and the r^{th} output variables.
- $b_{r,\, r+1}$ is the correlation coefficient between the r^{th} and $(r+1)^{\text{th}}$ output variables.

3. Comparison of various models with the CCRCOR model

3.1. The CCR model. The CCR model was first suggested by Charnes *et al.* [4] in 1978. The input oriented CCR model is formulated as follows:

$$
\begin{aligned}
(3.1)\qquad \max \theta_k = {} & \sum_{r=1}^{s} u_r y_{rk} \\
st.\quad & \sum_{i=1}^{m} v_i x_{ik} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0 \quad (j = 1, \ldots, n) \\
& u_1, u_2, \ldots, u_s \ge 0 \\
& v_1, v_2, \ldots, v_m \ge 0
\end{aligned}
$$

3.2. The CEM model. The CEM was proposed by Sexton *et al.* [14]. The purpose of the CEM is to rank the DMUs so as to select a DMU which has the best performance. The basic idea of this method is the use of a binary assessment rather than a self-evaluation in DEA.

The CEM's implementation phases are as follows:

(1) For each DMU_k, $(k = 1, 2, \ldots, n)$, a set of multipliers $w^*_{1k}, w^*_{2k}, \ldots, w^*_{mk}$, μ^*_{1k}, $\mu^*_{2k}, \ldots, \mu^*_{sk}$ is obtained.

(2) For any DMU_j, $(j = 1, 2, \dots, n)$, the cross-efficiency can be calculated as follows using the selected weights for DMU_k in the model (3.1).

$$E_{kj} = \frac{\sum_{r=1}^{s} \mu_{rk}^{*} y_{rj}}{\sum_{i=1}^{m} w_{ik}^{*} x_{ij}}, \quad k, j = 1, 2, \dots, n. \tag{3.2}$$

In the formulation (3.2), E_{kj} calculates the cross efficiency of DMU_j using the weights of DMU_k.

(3) As shown in Table 1, for each DMU, each of the columns of the cross-efficiency matrix is averaged to measure the mean cross-efficiency. Thus, for DMU_j, the cross-efficiency score is calculated as follows:

$$\bar{E}_j = \frac{1}{n} \sum_{k=1}^{n} E_{kj}. \tag{3.3}$$

Table 1. The generalized cross-efficiency matrix [22]

Evaluating DMU_k	The evaluated DMU_j				
	1	2	3	$\cdots$	n
1	E_{11}	E_{12}	E_{13}	$\cdots$	E_{1n}
2	E_{21}	E_{22}	E_{23}	$\cdots$	E_{2n}
3	E_{31}	E_{32}	E_{33}	$\cdots$	E_{3n}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
n	E_{n1}	E_{n2}	E_{n3}	$\cdots$	E_{nn}
mean	$\bar{E}_1$	$\bar{E}_2$	$\bar{E}_3$	$\cdots$	$\bar{E}_n$

The diagonal elements in Table 1 are the efficiency values which are calculated with the classical DEA (3.1).

The CEM method allows the elimination of a wrong classification for efficient DMUs. The CEM has three advantages [22]: First, it provides a ranking of the DMUs [14]. Second, the application area experts eliminate unrealistic weighting plans without the need for weight restrictions [2]. Third, the CEM can discriminate effectively between good and poor performance of the units [3]. Despite all these advantages, there are some disadvantages of the CEM. The most significant problem with the CEM is obtaining alternative optimal solutions for the weights. In this case, different cross-efficiency scores on the selected weight set are obtained [12]. It is recommended to use the second objectives in selecting the weights from the alternative solutions in the literature. The second purposes have been used widely in recent years. The basic idea of researchers who use this approach is to calculate an optimal set of weights for each DMU by installing some conditions on cross-efficiency. Benevolent and aggressive formulations are widely used on this issue [14, 7]. While the benevolent formulation maintains the efficiency score of the unit under evaluation, it will select the weights which will increase the cross efficiencies of the other DMUs. But while the aggressive formulation maintains the efficiency score of the unit under evaluation, it will select the weights which can reduce the cross efficiencies of the other DMUs [12]. Wu *et al.* [20] thought to maximize the sum of the ranking numbers for the DMUs except for the DMU under evaluation. They discuss the second objective functions which are defined with the principle of priority ranking. Wu *et al.* [20] put forward that tracing the best ranking is more important than maximizing the individual score. The purpose of Ramon *et al.* [12] is to avoid zero weights and large differences among the weights of the input and outputs. For this purpose, Ramon *et al.* [12] applied the CEM method without the resides as the expansion of Ramon *et al.* [11]. Wang and Chin's model [18] is called the neutral DEA model. The weights are determined without considering the impact on the other DMUs of each DMU in this model. In other words, the authors tried to reduce the number of zero weights in the outputs and to maximize the relative contribution of outputs with the min/max formulation in the CEM. Anderson *et al.* [2] emphasize that CEM should eliminate unrealistic weighting plans.

Due to alternative solutions obtained from the CEM, it can produce different cross efficiency scores. So, there is a lot of research in the literature about this issue. For example, the symmetric weight assignment technique (SWAT) is proposed as a second objective introduced into the CEM by Jahanshahloo *et al.* [8]. The SWAT provides the choice of weights in a symmetrical manner.

The average cross-efficiency is used widely in the literature. However, there are some disadvantages to the use of the average cross-efficiency as the final cross efficiency to evaluate and to rank the DMUs [22]: The average cross-efficiency measurement is not a Pareto solution. So, it is not good enough.

When the average of all cross-efficiencies is calculated, the relationships between weights will be lost [6]. Various authors have given up calculating the average of the cross-efficiency scores, in order to determine the final cross-efficiency scores. Instead, they used the principle of the priority ranking [20], namely the Nash equilibrium point [21], the Shapley value [19] and the Shannon entropy method [22]. The average cross efficiency assumption was abolished by Wu *et al.* [19].

The cross-efficiency is combined with game theory by Wu *et al.* [21]. Each candidate is considered as a player, who is trying to maximize his/her own efficiency in this approach. All of efficiencies are averaged to maximize the efficiency of each of the DMUs. Thus, the game cross-efficiency scores obtained constitute a Nash equilibrium point.

Liang *et al.* [10] focus on the alternative second purposes which are proposed as an alternative in the CEM. Wang and Chin [18] gave an application of new alternative models proposed by them. Wang and Chin [18] argued that they found more realistic results than those of the alternative models of Liang et al. [10].

3.3. The super efficiency model. The classical DEA assigns efficiency scores of 1 to the efficient units. On the other hand, for the inefficient units, it assigns efficiency scores of less than 1. Thus, the classical DEA can determine only the efficient units. But, DEA does not allow one to find the degree of efficiencies of the DMUs and to rank them.

Andersen and Petersen's Super Efficiency Method (AP) is the first ranking method, which was applied to compare and rank the efficient DMUs [1]. For this purpose, the evaluated DMU is excluded from the reference set. On the basis of ranking in the AP model, a DMU which has the highest efficiency score will be in the first order; but another DMU which has the lowest efficiency score, will be in the last order. In this way, according to the values of the super-efficiency score, all DMUs are sorted in descending order. Thus, the efficient DMUs are ranked.

An inefficient DMU in the CCR model is inefficient in the AP model. If the DMU under evaluation is inefficient, the inefficiency score of the DMU will have to be the same in both models. The AP model is formulated as follows:

$$\max \sum_{r=1}^{s} u_r y_{rk} \tag{3.4}$$
$$s.t. \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \ (j = 1, \ldots, n), \ j \neq k$$
$$\sum_{i=1}^{m} v_i x_{ik} = 1$$
$$u_r, v_i \geq \varepsilon, \ r = 1, \ldots, s; \ i = 1, \ldots, m$$

3.4. True efficiency ratio. The true efficiency ratio is calculated as the ratio of the sum of the weighted outputs to the sum of the weighted inputs [9]. This ratio, which was used by Thanassoulis [16], Lam [9], etc., will be examined in detail with applications in the next section.

4. Applications

4.1. Example. In this study, we applied the suggested CCRCOR model and other methods to the data set (Table 2) which was used by Sherman [15], Thanassoulis [16] and Lam [9], etc. The efficiencies of 15 hypothetical hospitals were evaluated by the data set. The set of data is composed from 3 outputs (severe patients (SP), regular patients (RP), the number of teaching units (TU)) and an input (total cost). The equality (4.1), which is established for the set of data in Sherman [15], has been used in the studies by many authors (Thanassoulis[16] and Lam [9], etc.):

$$\text{Total Cost} = 0.5\text{TU} + 133.68\text{RP} + 174.74\text{SP} \tag{4.1}$$

Table 2. 15 hospitals with a single input and three outputs (Sherman [15] **and Thanassoulis** [16]**)**

Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost (v_1) Input 1 ($)
1	50	3	2	775.5
2	50	2	3	816.6
3	100	2	3	841.6
4	100	3	2	800.5
5	50	3	3	950.3
6	100	2	5	1191.05
7	50	10	2	1711.3
8	100	3	2	884.75
9	50	2	3	841.6
10	100	10	2	2036.3
11	50	5	3	1362.6
12	100	3	3	1070
13	50	4	5	1491.1
14	50	3	2	898.7
15	100	3	3	1070

When the set of data in Table 2 is analyzed, the correlation coefficients between many inputs and outputs are at unexpectedly low levels as seen in Table 3. Significant relationships between inputs and outputs in a production cluster are expected. The set of data in Table 2, which was used by Sherman [15], Thanassoulis [16] and Lam [9], should have been initially analyzed in this respect. Therefore, considering that the results are obtained with this unsuitable data set, the CCRCOR results (Table 4) are quite good.

Meaningless relationships between inputs and outputs can be seen in Table 2, even with the naked eye without creating the correlation matrix. Relationships between variables in a production process should not be contradictory. In the CCRCOR model we make the assumption that there should be a significant correlation between inputs and outputs in a production relationship. Nevertheless, the results can be considered quite good. Instead of the set of data used here, our suggestion is to use real production data sets or data sets which are produced by simulation with correlations above a certain level between inputs and outputs. So, we think that working with such data types can allow more meaningful assessments.

Table 3. The correlation matrix between a single input and three outputs for 15 hospitals in Example 4.1

		Cost	TU	RP	SP
Correlation	Cost	1.000	.030	.886	.093
	TU	.030	1.000	-.056	-.009
	RP	.886	-.056	1.000	-.338
	SP	.093	-.009	-.338	1.000
Sig. (1-tailed)	Cost		.458	.000	.370
	TU	.458		.421	.487
	RP	.000	.421		.109
	SP	.370	.487	.109	

According to the correlation matrix in Table 3, the relationship between the RP and the Cost (0.886) is significant. But the relationships between the other variables are not significant in Table 3. If the model can be explained with only one variable, it is meaningless adding other variables into the model (the equality (4.1)). In regression analysis, there should be a significant linear correlation between the dependent and independent variables. In the theory of linear regression, variables which have got insignificant correlations should not be included in the model. So, the data set cannot trusted. Although this is a hypothetical data set, it must fit roughly with production theory and regression theory. The relations between the input and output variables of the data set are meaningless. So, there are low and reverse correlations between the input and output variables. In order to compare the CCRCOR model with other methods, the analysis was continued because we were obliged to work with this data set. In Table 4, the first seven hospitals are found efficient in the evaluation by the true efficiency ratio and the CCR model. According to the CCRCOR model results, only the fourth DMU is efficient.

In Example 4.1, the true efficiency ratio and the CCR model had the Pearson correlation (Spearman's rho) equal to 0.682 (0.722) and 0.613 (0.617), respectively; while the CCRCOR model obtained a good performance with the Pearson correlation and Spearman's rho equal to 0.778 and 0.796, respectively (see Table 4). According to the Super Efficiency model, Pearson correlation and Spearman's rho are equal to 0.613 and 0.553, respectively. In Table 5, while the number of zero weights which are obtained by the CCR model is 11; the number of zero weights which are obtained by the CCRCOR model is 4. Thus, the number of zero weights using the CCRCOR model is greatly reduced.

Table 4. The assessments with a single input for 15 hospitals in Example 4.1

Hospital	CCRCE	Trueeffc	CCR	Supereffc	CCRCOR
1	0.82389 (4)*	1 (4)	1 (4)	1.00002 (6)	0.53217 (7)
2	0.81143 (6)	1 (4)	1 (4)	1.00400 (4)	0.51717 (8)
3	0.89121 (2)	1 (4)	1 (4)	1.10585 (3)	0.96394 (2)
4	0.96735 (1)	1 (4)	1 (4)	1.10525 (7)	1 (1)
5	0.80335 (7)	1 (4)	1 (4)	1.00099 (5)	0.44559 (11)
6	0.83881 (3)	1 (4)	1 (4)	1.14969 (2)	0.69918 (6)
7	0.78576 (11)	1 (4)	1 (4)	1.18991 (1)	0.24578 (15)
8	0.82103 (5)	0.90480 (12)	0.90478 (13)	0.90478 (13)	0.90478 (3)
9	0.78733 (10)	0.97027 (8)	0.97030 (9)	0.97030 (9)	0.50181 (9)
10	0.70337 (15)	0.85266 (15)	0.92334 (10)	0.92334 (10)	0.39700 (12)
11	0.70833 (14)	0.89360 (13)	0.89516 (14)	0.89516 (14)	0.31242 (13)
12	0.79528 (9)	0.91145 (10.5)	0.91144 (11.5)	0.91144 (11.5)	0.75818 (4.5)
13	0.74668 (12)	0.96132 (9)	0.97253 (8)	0.97253 (8)	0.29916 (14)
14	0.71102 (13)	0.86294 (14)	0.86291 (15)	0.86291 (15)	0.45921 (10)
15	0.79528 (8)	0.91146 (10.5)	0.91144 (11.5)	0.91144 (11.5)	0.75818 (4.5)
Pearson correlation**	1	0.682 (p=0.003)	0.613 (p=0.008)	0.613 (p=0.008)	0.778 (p=0.000)
Spearman's rho	1	0.722 (p=0.001)	0.617 (p=0.007)	0.553 (p=0.016)	0.796 (p=0.000)

*Efficiency ranking of 15 hospitals
**Pearson correlation and Spearman's rho with CCRCE

The true efficiency ratio seen in Table 4 is calculated with the following formulation:

"True efficiency ratio = estimated cost (for Equality (4.1)) / actual cost".

Table 5. The weights obtained with the CCRCOR model and the CCR model for a single input

CCRCOR	Weights			
Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost (v_1) Input 1
1	0.010001	0.001455	0.013866	0.001289
2	0.009498	0.001382	0.013168	0.001225
3	0.009256	0	0.012776	0.001188
4	0.009700	0	0.013400	0.001200
5	0.008161	0.001188	0.011315	0.001052

Table 5. Continued

CCRCOR	Weights			
Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost (v_1) Input 1
6	0.006540	0	0.009028	0.000840
7	0.004532	0.000660	0.006283	0.000584
8	0.008766	0.001276	0.012153	0.001130
9	0.009216	0.001341	0.012776	0.001188
10	0.003809	0.000554	0.005281	0.000491
11	0.005692	0.000828	0.007891	0.000734
12	0.007280	0	0.010049	0.000935
13	0.005202	0.000757	0.007211	0.000671
14	0.008630	0.001256	0.011965	0.001113
15	0.007249	0.001055	0.010049	0.000935
CCR	Weights			
Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost (v_1) Input 1
1	0.000645	0.172386	0.225302	0.001289
2	0.000609	0.163707	0.214038	0.001225
3	0.008535	0	0.048836	0.001188
4	0.001000	0	0	0.001200
5	0.000516	0.140692	0.184047	0.001052
6	0	0	0.200000	0.000840
7	0.001974	0.090131	0	0.000584
8	0.009048	0	0	0.001130
9	0.000591	0.158844	0.207680	0.001188
10	0.001659	0.075746	0	0.000491
11	0	0.09887	0.133602	0.000734
12	0.000467	0.124928	0.163309	0.000935
13	0	0.089665	0.122773	0.000671
14	0.000556	0.148755	0.194416	0.001113
15	0.000467	0.124928	0.163309	0.000935

4.2. Example. The original data set given in Table 2 has only one input. DEA is usually used to evaluate multiple inputs and outputs. So, the cost, which is an input of the original data set, is classified into two elements using Equation (4.2) defined by Lam[9]:

$$\text{Total Cost} = 25\text{Cost1} + 50\text{Cost2} \tag{4.2}$$

The original total cost of each hospital remains the same after this classification. Equation (4.3) is obtained from Equation (4.1) and Equation (4.2):

$$25\text{Cost1} + 50\text{Cost2} = 0.5\text{TU} + 133.68\text{RP} + 174.74\text{SP} \tag{4.3}$$

The set of data given in Table 6 was obtained in this way by Lam [9]. The set of data with 15 DMUs consists from two inputs (Cost 1, Cost 2) and three outputs (TU, RP and SP). DMUs $1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 13$ were found efficient by the CCR model. By the True Efficiency, the first 7 hospitals were efficient. DMUs $3, 4, 8, 12$ were efficient with the CCRCOR model.

Table 6. The inputs and outputs of 15 hospitals with two inputs in Example 4.2 (Lam [9], page 139 (for the first 15 DMUs))

Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost 1 (v_1) Input 1	Cost 2 (v_2) Input 2
1	50	3	2	11	10.01
2	50	2	3	22	5.332
3	100	2	3	16	8.832
4	100	3	2	15	8.51
5	50	3	3	27	5.506
6	100	2	5	31	8.321
7	50	10	2	28	20.226
8	100	3	2	23	6.195
9	50	2	3	19	7.332
10	100	10	2	30	25.726
11	50	5	3	21	16.752
12	100	3	3	13	14.9
13	50	4	5	24	17.822
14	50	3	2	16	9.974
15	100	3	3	25	8.9

Table 7. The correlation matrix between two inputs and three outputs for the 15 hospitals in Example 4.2

		Cost 1	Cost 2	TU	RP	SP
Correlation	Cost 1	1.000	.278	.071	.462	.344
	Cost 2	.278	1.000	.001	.866	-.061
	TU	.071	.001	1.000	-.056	-.009
	RP	.462	.866	-.056	1.000	-.338
	SP	.344	-.061	-.009	-.338	1.000

Table 7 (continued)

		Cost 1	Cost 2	TU	RP	SP
Sig. (1-tailed)	Cost 1		.158	.401	.041	.105
	Cost 2	.158		.499	.000	.415
	TU	.401	.499		.421	.487
	RP	.041	.000	.421		.109
	SP	.105	.415	.487	.109	

According to Table 7, while the relations between Cost 1 and Cost 2 with RP are meaningful, the relations between other variables are meaningless. The correlations between the inputs (Cost 1, Cost 2) and only the RP are significant for the data sets in Example 4.2. In this case, it is meaningless to include other variables in the model.

Table 8. The assessments for 15 hospitals with two inputs in Example 4.2

Hospital	CCRCE	Trueeffc	CCR	Supereffc	CCRCOR
1	0.86716 (4)*	1 (4)	1 (6)	1.05632 (8)	0.79243 (7)
2	0.78701 (8)	1 (4)	1 (6)	1.00369 (11)	0.71984 (8)
3	0.95809 (1)	1 (4)	1 (6)	1.14963 (7)	1 (2.5)
4	0.95626 (2)	1 (4)	1 (6)	1.21556 (5)	1 (2.5)
5	0.75579 (11)	1 (4)	1 (6)	1.25612 (3)	0.67927 (11)
6	0.84688 (5)	1 (4)	1 (6)	1.17977 (6)	0.87941 (5)
7	0.73608 (13)	1 (4)	1 (6)	1.24278 (4)	0.55803 (15)
8	0.81996 (6)	0.90480 (12)	1 (6)	1.37369 (1)	1 (2.5)
9	0.78325 (9)	0.97027 (8)	0.97737 (12)	0.97737 (12)	0.68688 (10)
10	0.73998 (12)	0.85266 (15)	1 (6)	1.02353 (9)	0.58939 (13)
11	0.71063 (15)	0.89360 (13)	0.90179 (14)	0.90179 (14)	0.57246 (14)
12	0.91301 (3)	0.91145 (10.5)	1 (6)	1.30769 (2)	1 (2.5)
13	0.76139 (10)	0.96132 (9)	1 (6)	1.03779 (10)	0.69476 (9)
14	0.71575 (14)	0.86294 (14)	0.86293 (15)	0.86293 (15)	0.62806 (12)
15	0.80372 (7)	0.91145 (10.5)	0.92573 (13)	0.92573 (13)	0.84861 (6)
Pearson correlation**	1	0.426 (p=0.057)	0.472 (p=0.038)	0.480 (p=0.035)	0.896 (p=0.000)
Spearman's rho	1	0.475 (p=0.037)	0.500 (p=0.029)	0.475 (p=0.037)	0.933 (p=0.000)

Efficiency ranking of 15 hospital

**Pearson correlation and Spearman's rho with CCRCE

In Table 8, the maximum correlation is between the CCRCOR and CCRCE models. There the Pearson's correlation (Spearman's rho) is equal to 0.896 (0.933) between the CCRCOR and the CCRCE models. In this case, the CCRCOR model is the closest to the CCRCE model in efficiency ranking. Thus, the CCRCOR model is suggested as an alternative to the CCRCE model.

Table 9. The weights obtained from the CCRCOR and the CCR with two inputs in Example 4.2

CCRCOR	Weights				
Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost 1 (v_1) Input 1	Cost 2 (v_2) Input 2
1	0.00658	0.04120	0.16988	0.05844	0.03568
2	0.00794	0.05354	0.07191	0.02474	0.08548
3	0.00667	0.03296	0.08889	0.04675	0.02854
4	0.00707	0.03491	0.09416	0.04951	0.03023
5	0.00700	0.04703	0.06316	0.02173	0.07508
6	0.00627	0.00692	0.04771	0.01641	0.05904
7	0.00283	0.03284	0.04410	0.01517	0.02844
8	0.00723	0.04874	0.06546	0.02252	0.07782
9	0.00552	0.05238	0.10211	0.03513	0.04536
10	0.00313	0.02767	0	0.01278	0.02396
11	0.00361	0.02258	0.09309	0.03202	0.01955
12	0.00646	0.03191	0.08606	0.04525	0.02763
13	0.00178	0	0.12112	0.04167	0
14	0.00544	0.05163	0.10066	0.03463	0.04471
15	0.00650	0.01425	0.05214	0.01794	0.06198
CCR	Weights				
Hospital	TU (u_1) Output 1	RP (u_2) Output 2	SP (u_3) Output 3	Cost 1 (v_1) Input 1	Cost 2 (v_2) Input 2
1	0.00343	0.19271	0,12524	0.05688	0.03740
2	0.00061	0.16372	0.21401	0.03061	0.06124
3	0.00829	0	0.05692	0.04910	0.02428
4	0.01000	0	0	0.05661	0.01772
5	0.00053	0.14069	0.18388	0.02631	0.05263
6	0.00042	0.11221	0.14671	0.02099	0.04197
7	0.00143	0.08929	0.01786	0.03571	0
8	0.00905	0	0.04758	0.02252	0.07782
9	0	0.15418	0.22301	0.03057	0.05717
10	0.00905	0	0.04758	0.02252	0.07782
11	0	0.09840	0.13660	0.01883	0.03610
12	0.00049	0.09789	0.21902	0.03730	0.03457
13	0	0	0.20000	0.03341	0.01112
14	0.00056	0.14875	0.19443	0.02782	0.05564
15	0.00200	0.11251	0.12925	0.01794	0.06198

According to Table 9, while the number of zero weights obtained by the CCR model is 10, the number of zero weights obtained by the CCRCOR is 3. So, the number of zero weights found by the CCRCOR decreased remarkably. In addition, the value of the relationship between the CCRCE model and the True efficiency is calculated as 0.475 (see Table 8). The results which are obtained from the CCRCOR model which takes into account the correlations between input and output variables should be regarded as satisfactory.

5. Conclusion

We recommend a new model (the CCRCOR) which adds to the CCR model by taking as weight constraints the relations among the variables given by the correlation coefficients. In the DEA literature, authors engaged in weight balancing have worked without taking into account the relationships between input and output variables. In previous studies, the correlations between the variables were excluded from consideration prior to the analysis. For this reason, variables which have a meaningless or minimal relationship, or which are related in the opposite direction, have participated in the analysis unbeknown to the authors. This situation causes unreasonable results. Therefore, in this paper we have created the CCRCOR model with regulations in the CCR at the level of relationships between input and output variables. The CCRCOR, which is an objective model, does not require prior knowledge or value judgments, unlike other weight restriction methods. The CCRCE model can calculate different efficiency scores because of alternative solutions. CCRCOR reduces the chance of finding an alternative solution because it restricts the set of weights according to the cross efficiency evaluation method. But, as with other methods of weight restriction, sometimes an optimal solution in CCRCOR may not be available. If the optimal solution does not exist, the weight restrictions should be relaxed. In addition, the CCRCOR model takes account the relationships between variables. So, the CCRCOR model eliminates the problem of weight flexibility in the CCR model. Also, the CCRCOR model produces a smaller number of zero weights compared to the CCR model.

Though the data sets used in the applications presented in the two examples in the previous section are not trustworthy, the CCRCOR model gave quite good results. Also, the similarity in the ranking between the CCRCE model and the CCRCOR model is striking. If it had been used with a more realistic data set, or an artificial data set derived from simulation taking into account the relationships between variables, instead of these data sets frequently used in the literature (Thanassoulis [16] and Lam [9]), the results could have been even more reliable.

References

[1] Andersen, P. and Petersen, N. C. *A procedure for ranking efficient units in data envelopment analysis*, Management Science **39**, 1261–1264, 1993.

[2] Anderson, T. R., Hollingsworth, K. and Inman, L. *The fixed weighting nature of a cross-evaluation model*, Journal of Productivity Analysis **17**, 249–255, 2002.

[3] Boussofiane, A., Dyson, R. G. and Thanassoulis, E. *Applied data envelopment analysis*, European Journal of Operational Research **52**, 1–15, 1991.

[4] Charnes, A., Cooper, W. W. and Rhodes, E. *Measuring the efficiency of decision making units*, European Journal of Operational Research **2**, 429–444, 1978.

[5] Cooper, W. W., Seiford, L. M. and Tone, K. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, second ed., (Springer Science & Business Publishers, New York, 2007).

[6] Despotis, D. K. *Improving the discriminating power of DEA: Focus on globally efficient units*, Journal of the Operational Research Society **53**, 314–323, 2002.

[7] Doyle, J. and Green, R. *Efficiency and cross-efficiency in DEA; Derivations, meanings and uses*, Journal of The Operational Research Society **45**, 567–578, 1994.

[8] Jahanshahloo, G. R., Hosseinzadeh Lotfi, F., Jafari, Y. and Maddahi, R. *Selecting symmetric weights as a secondary goal in DEA cross-efficiency evaluation*, Applied Mathematical Modelling **35**, 544–549, 2011.

[9] Lam, K. F. *In the determination of weight sets to compute cross-efficiency ratios in DEA*, Journal of the Operational Research Society **61**, 134–143, 2010.

[10] Liang, L., Wu, J. , Cook, W. and Zhu, J. *Alternative secondary goals in DEA cross efficiency evaluation*, Int. J. Production Economics **113**, 1025–1030, 2008.

[11] Ramón, N., Ruiz, J. L. and Sirvent, I. *A multiplier bound approach to assess relative efficiency in DEA without slacks*, European Journal of Operational Research **203**, 261–269, 2010.

[12] Ramón, N., Ruiz, J. L. and Sirvent, I. *On the choice of weights profiles in cross-efficiency evaluations*, European Journal of Operational Research **207**, 1564–1572, 2010.

[13] Roll, Y. and Golany, B. *Alternate methods of treating factor weights in DEA*, Omega **21**, 99–109, 1993.

[14] Sexton, T. R., Silkman, R. H. and Hogan, A. J. *Data envelopment analysis: Critique and extensions*, in R. H Silkman (Ed.), *Measuring Efficiency: An assessment of Data Envelopment Analysis, Jossey-Bass, San Francisco*, **32**, 73–105, 1986.

[15] Sherman, H. D. *Data envelopment analysis as a new managerial audit methodology- test and evaluation*, Auditing: A Journal of Practice and Theory **4**, 35–53, 1984.

[16] Thanassoulis, E. *A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments*, Journal of the Operational Research Society **44**, 1129–1144, 1993.

[17] Thompson, R. G., Singleton, F. D., Thrall, R. M. and Smith, B. A. *Comparative site evaluations for locating a high-energy physics lab in Texas*, Interfaces **16**, 35–49, 1986.

[18] Wang, Y. M. and Chin, K. S. *A neutral DEA model for cross-efficiency evaluation and its extension*, Expert Systems with Applications **37**, 3666–3675, 2010.

[19] Wu, J., Liang, L. and Yang, F. *Determination of the weights for the ultimate cross efficiency using Shapley value in cooperative game*, Expert Systems with Applications **36**, 872–876, 2009.

[20] Wu, J., Liang, L. and Yang, F. *Determination of cross-efficiency under the principle of ranking priority in cross-evaluation*, Expert Systems with Applications **36**, 4826–4829, 2009.

[21] Wu, J., Liang, L. and Zha, Y. *Preference voting and ranking using DEA game cross efficiency model*, Journal of the Operations Research Society of Japan **52**, 105–111, 2009.

[22] Wu, J., Sun, J., Liang, L. and Zha, Y. *Determination of weights for ultimate cross efficiency using Shannon entropy*, Expert Systems with Applications **38**, 5162–5165, 2011.