

# Using Cross-Evaluations in Data Envelopment Analysis

Kim Fung Lam

**Abstract**—In data envelopment analysis (DEA), all efficient decision making units (DMUs) have the same efficiency score which are difficult to discriminate. Cross-efficiency can be used to discriminate among efficient-DMUs. It also provides an alternative efficiency ranking for inefficient-DMUs. However, weight sets used for cross-efficiency may not be proper for ranking DMUs. The reason for this is that alternate optimal solutions for the weights exist for efficient-DMUs in DEA. The weight sets for efficient-DMUs are usually picked arbitrarily by linear programming (LP) software within the alternate optimal region. It is highly unlikely that those weight sets obtained from the LP software are the most suitable for performance measure. Therefore, some researchers have proposed using secondary objectives to search for more homogeneous or better weight sets within the alternate optimal regions. This paper examines four of those models and attempts to determine the accuracy of the four models in estimating the performance rankings for DMUs in DEA.

**Index Terms**—Cross-efficiency, DEA, efficiency score, performance measure.

## I. INTRODUCTION

Consider  $n$  decision making units,  $DMU_j$  ( $j = 1, 2, \dots, n$ ) each of which has  $m$  inputs  $x_{ij}$  ( $i = 1, \dots, m$ ) and  $s$  outputs  $y_{rj}$  ( $r = 1, \dots, s$ ). In DEA, DMUs are classified into either efficient set ( $E$ ) or inefficient set ( $\bar{E}$ ). The relative efficiency of each  $DMU_0$  can be obtained from the following linear programming (LP) model [1]:

$$\text{Max } \sum_{r=1}^s u_r y_{r0} \quad (1)$$

$$\text{s. t. } \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, \quad j = 1, \dots, n, \quad (2)$$

$$\sum_{i=1}^m v_i x_{i0} = 1, \quad (3)$$

where  $u_r, v_i \geq 0$ ,  $r = 1, \dots, s$ ;  $i = 1, \dots, m$ , are factor weights for output  $r$  and input  $i$ , respectively. Another version of the model requires  $u_r, v_i > 0$ .

One common problem in DEA is that all efficient-DMUs have the same efficiency score which makes it difficult to discriminate among efficient-DMUs. Many methods have been proposed to rank DMUs, especially efficient-DMUs in DEA. A comprehensive review of ranking methods in the DEA context is given in [2]. Among the different methods, Cross-efficiency [3] is one of the most popular methods for ranking DMUs. Cross-efficiency score of a DMU is the average of efficiency scores computed by using input weights and output weights obtained from all DMUs. However, it is well known that alternate optimal solutions of the weights

exist for the LP solutions of efficient-DMUs. Which optimal weight set is chosen usually depends on which optimal solution is first found by the LP software. Different software used may result in choosing different weight sets as pointed out by [4]. This ample choice of weights in the alternate optimal regions makes the first optimal solution found by the software not necessarily a good choice for performance measure.

Some secondary objectives [3]-[8] have been proposed to search for more homogeneous or better weight sets within the alternate optimal regions. Among them [3], [5] proposed four LP models using different secondary objectives to determine weight sets to compute cross-efficiency scores. However, among the four models, which model can provide a more accurate performance ranking of DMUs and also how good those models when compared to simple cross-efficiency are have not been further discussed in the literature. In view of this, this paper discusses characteristics of these models and also examines their performance via a simulation experiment. The main objective of this paper is to find out the performance of the four models and how good those models are when compared with simple cross-efficiency in providing a more accurate performance ranking of DMUs.

## II. THE FOUR MODELS

Instead of using the original weight sets obtained directly from DEA, [3], [5] proposed several LP models, each uses a secondary objective to determine a new weight set for each DMU. We examine four of those models in this paper. In general, each of the four models can be divided into two stages. Stage one is to obtain simple efficiency,  $e_0$  for each  $DMU_0$  using DEA. The CCR model [1] can be employed to obtain efficiencies for DMUs in stage one. In stage two, for each decision making unit,  $DMU_0$ , a secondary objective is applied to search for a new weight set for each DMU. The stage two formulations of the four models are listed in the following paragraphs.

The Benevolent I (BE-I) model is proposed by [3] and is stated as follows:

$$\text{Max } \sum_{j=1}^n \left( \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \right)_{j \neq 0} \quad (4)$$

$$\text{s. t. } \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, \quad j = 1, \dots, n, \quad (5)$$

$$\sum_{i=1}^m v_i x_{i0} = 1, \quad (6)$$

$$\sum_{r=1}^s u_r y_{r0} - e_0 \sum_{i=1}^m v_i x_{i0} = 0, \quad (7)$$

where  $u_r, v_i \geq 0$ ,  $r = 1, \dots, s$ ;  $i = 1, \dots, m$ ;  $e_0$  is the efficiency score of  $DMU_0$  determined in stage 1. Alternatively, the LP formulation of Aggressive I (AG-I) can be obtained by changing the objective function of BE-I from maximizing to

Manuscript received April 24, 2014; revised June 27, 2014.

Kim Fung Lam is with the Department of Management Sciences, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong (e-mail: msblam@cityu.edu.hk).

minimizing.

The Benevolent II (BE-II) model is proposed by [5] and is stated as follows:

$$\text{Max } \sum_{j=1}^n (\sum_{r=1}^s u_r y_{rj}) \quad (8)$$

$$\text{s. t. } \sum_{j=1}^n (\sum_{i=1}^m v_i x_{ij}) = 1, \quad (9)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, \quad j = 1, \dots, n, \quad (10)$$

$$\sum_{r=1}^s u_r y_{r0} - e_o \sum_{i=1}^m v_i x_{i0} = 0, \quad (11)$$

where  $u_r, v_i \geq 0, r = 1, \dots, s; i = 1, \dots, m; e_o$  is the efficiency score of  $DMU_o$  determined in stage 1. Alternatively, the LP formulation of Aggressive II (AG-II) can be obtained by changing the objective of BE-II from maximizing to minimizing.

### III. MODEL DISCUSSION

If the redistribution of resources is based on the efficiencies of the DMUs, then the actual resource redistribution among the DMUs may be significantly affected by which secondary objective are being applied. BE-I and BE-II maximize the wellbeing of all DMUs. Their performance measures tend to give higher efficiencies to most of the DMUs when compared with the aggressive models. These objectives may have certain important economic interpretation in terms of DMUs' performance and also input and output weights structures, however, it is not necessary that the underlying weight structure should maximize the wellbeing of all DMUs.

The objectives of benevolent models are to maximize the wellbeing of all DMUs, then it is expected that the chosen weight sets should provide higher efficiency scores for most DMUs. In general, those weight sets are expected to give more balanced virtual inputs and outputs since extremely unequal virtual inputs and outputs may tend to lower efficiency scores for most DMUs. As a result, extremely unequal virtual inputs and outputs are expected to be less common in both BE-I and BE-II. As pointed out by [4], it is better to use weight sets which have more balanced virtual inputs and outputs than extreme virtual inputs and outputs in DEA. This is shown in their model which seeks to maximize the minimum virtual variables as one of the optimizing objectives. Extreme virtual inputs and outputs sometime may put most weights on a small number of inputs or outputs. Some inputs or outputs may receive zero or close to zero virtual values. Using it as an instrument to measure efficiency of DMUs is not that appropriate because it may be difficult to explain why such an extreme virtual inputs and outputs structure exists in a rational economy. Both AG-I and AG-II have the tendency to obtain weight sets which produce more unequal virtual inputs and outputs. In the next section, we conduct a simulation experiment to study the performance of the four methods as well as the simple cross-efficiency method (SC-E) in the rankings of DMUs.

### IV. DESIGN OF THE SIMULATION EXPERIMENT

In this simulation experiment, we assume that the

production function follows a simple linear form. We make use of the linear production function to generate efficient and inefficient DMUs. Then we apply DEA, SC-E and the four models to the generated data sets to obtain efficiency rankings. The efficiency rankings obtained from all the methods are compared with the efficiency ranking generated from the original linear production function. Details of our simulation experiment are given as follows.

In the simulation experiment, we assume that the production function takes the following linear form:

$$\sum_{i=1}^m v_i x_{ij} = \sum_{r=1}^s u_r y_{rj}. \quad (12)$$

In this experiment, we use three inputs, three outputs, and three different weight sets, namely, as follows:

$$\begin{aligned} S_1 &= \{v_1=1; v_2=1; v_3=1; u_1=1; u_2=1; u_3=1\}, \\ S_2 &= \{v_1=1; v_2=1; v_3=9; u_1=1; u_2=1; u_3=9\}, \\ S_3 &= \{v_1=1; v_2=9; v_3=9; u_1=1; u_2=9; u_3=9\}. \end{aligned}$$

The three weight sets:  $S_1$  represents no dominant input or output scenario;  $S_2$  represents one dominant input and one dominant output scenario;  $S_3$  represents one minor input and one minor output scenario.

TABLE I: LIST OF 27 CASES IN THE SIMULATION EXPERIMENT

Case	Sample Size	Weight Set $\{v_1 v_2 v_3\} \{u_1 u_2 u_3\}$	Combination of DMUs in the three Efficiency Classes: (A E I)
1	15	{1 1 1} {1 1 1}	(0.25 0.25 0.50)
2	15	{1 1 1} {1 1 1}	(0.50 0.00 0.50)
3	15	{1 1 1} {1 1 1}	(0.00 0.50 0.50)
4	15	{1 1 9} {1 1 9}	(0.25 0.25 0.50)
5	15	{1 1 9} {1 1 9}	(0.50 0.00 0.50)
6	15	{1 1 9} {1 1 9}	(0.00 0.50 0.50)
7	15	{1 9 9} {1 9 9}	(0.25 0.25 0.50)
8	15	{1 9 9} {1 9 9}	(0.50 0.00 0.50)
9	15	{1 9 9} {1 9 9}	(0.00 0.50 0.50)
10	30	{1 1 1} {1 1 1}	(0.25 0.25 0.50)
11	30	{1 1 1} {1 1 1}	(0.50 0.00 0.50)
12	30	{1 1 1} {1 1 1}	(0.00 0.50 0.50)
13	30	{1 1 9} {1 1 9}	(0.25 0.25 0.50)
14	30	{1 1 9} {1 1 9}	(0.50 0.00 0.50)
15	30	{1 1 9} {1 1 9}	(0.00 0.50 0.50)
16	30	{1 9 9} {1 9 9}	(0.25 0.25 0.50)
17	30	{1 9 9} {1 9 9}	(0.50 0.00 0.50)
18	30	{1 9 9} {1 9 9}	(0.00 0.50 0.50)
19	60	{1 1 1} {1 1 1}	(0.25 0.25 0.50)
20	60	{1 1 1} {1 1 1}	(0.50 0.00 0.50)
21	60	{1 1 1} {1 1 1}	(0.00 0.50 0.50)
22	60	{1 1 9} {1 1 9}	(0.25 0.25 0.50)
23	60	{1 1 9} {1 1 9}	(0.50 0.00 0.50)
24	60	{1 1 9} {1 1 9}	(0.00 0.50 0.50)
25	60	{1 9 9} {1 9 9}	(0.25 0.25 0.50)
26	60	{1 9 9} {1 9 9}	(0.50 0.00 0.50)
27	60	{1 9 9} {1 9 9}	(0.00 0.50 0.50)

In each of the 27 cases, we generate 500 data sets of input and output scores according to the experimental design in Table I. For each data set, we obtain the simple efficiency score of each DMU through DEA. We define this efficiency score as  $e_o$  for  $DMU_o$ . Then, for  $DMU_o$ , we solve the four models in turns using  $e_o$ . We repeat the step for all DMUs. We then obtain a weight set for each DMU for each model. We use the weight sets obtained from each model to compute cross-efficiency scores for all DMUs. Intuitively, if the weight set obtained from a model is close to the underlying

weight set, then the performance rankings of DMUs obtained from the model and the underlying weight set should be similar to each other. Consequently, we compute Pearson correlation coefficients and also Spearman's rho coefficients between the true efficiency scores and the cross-efficiency scores obtained from each model. For comparison purpose, we also compute correlations between the true efficiency scores and the efficiency scores from DEA, and also the efficiency scores from SC-E.

In our simulated samples, all input scores and output scores of all DMUs are generated from a uniform distribution. We generate three different classes of DMUs according to their productive efficiencies, namely, Advanced-DMU (A), Efficient-DMU (E), and Inefficient-DMU (I). For an Efficient-DMU, the sum of weighted inputs is equal to the sum of weighted outputs. For an Inefficient-DMU, we randomly add more inputs to and subtract some amount of outputs from an Efficient-DMU. As a result, for an Inefficient-DMU, the sum of weighted inputs is more than the sum of weighted outputs. Inefficient-DMUs are usually weaker DMUs and their performance is very often below the industrial average. For an Advanced-DMU, we randomly subtract some inputs from and add more outputs to an Efficient-DMU. Hence, for an Advanced-DMU, the sum of weighted inputs is less than the sum of weighted outputs. In some cases of this simulation experiment, we include Advanced-DMUs to reflect the situation that sometimes a few more pioneering DMUs may always outperform other DMUs which have average performance in the industry. In some other cases, we only generate Efficient-DMUs and Inefficient-DMUs. In those cases, the true efficiency scores are more similar to DEA efficiency scores, i.e., efficiency scores of efficient and inefficient-DMUs are equal to one and less than one, respectively. We use three different sample sizes, namely, 15, 30 and 60 in this simulation experiment. Using three different sample sizes, three different weight sets, and three different mixes of DMU efficiency classes, we can obtain 27 cases or scenarios. The layouts of the 27 cases are summarized in Table I.

V. RESULTS OF THE SIMULATION EXPERIMENT

Pearson correlation coefficients and Spearman's rho coefficients are reported in Table II and Table II, respectively.

In Table II and Table III, the correlation coefficients of SC-E and the true efficiency are much higher than those of DEA and the true efficiency, and the differences are all statistically significant. This result has confirmed the value of applying cross-efficiency in DEA. When comparing SC-E with the other four models, SC-E has correlation coefficients less than both BE-I and BE-II in general and also those differences are statistically significant in most cases, but correlation coefficients of SC-E are all higher than AG-I and AG-II. These results may be explained by our previous discussions about the characteristics of the four models. Since both benevolent models maximize the wellbeing of all DMUs, it is expected that they are less likely to choose extreme virtual inputs and outputs when compared to the two aggressive models. Aggressive models are expected to choose extreme virtual inputs and outputs more often in order to minimize overall efficiencies of other DMUs. Intuitively,

extreme virtual inputs and outputs can provide strange results in reality since the weights obtained may be unrealistic and are difficult to explain in a rational economy.

TABLE II: MEAN PEARSON CORRELATION COEFFICIENTS AND STANDARD DEVIATIONS BETWEEN THE TRUE EFFICIENCY SCORES AND THE EFFICIENCY SCORES OF THE SIX METHODS

Mean Pearson Correlation Coefficients: the true efficiency scores with efficiency scores of method- <i>i</i>						
Case	DEA	SC-E	BE-I	BE-II	AG-I	AG-II
1	0.4104	0.8005	0.8030	0.8024	0.6330	0.6344
	0.1417	0.1404	0.1352	0.1358	0.2019	0.2002
2	0.5357	0.8558	0.8642	0.8631	0.6896	0.6893
	0.1412	0.1212	0.1177	0.1182	0.1885	0.1883
3	0.4158	0.8248	0.84420.	0.84270.	0.4521	0.4547
	0.1725	0.1454	1376	1404	0.2528	0.2515
4	0.4822	0.7726	0.7865	0.7925	0.7271	0.7269
	0.1699	0.1420	<sup>b</sup> 0.1378	<sup>b</sup> 0.1377	0.1439	0.1453
5	0.6013	0.8057	0.8245	0.8280	0.7543	0.7532
	0.1702	0.1493	<sup>b</sup> 0.1394	<sup>b</sup> 0.1377	0.1439	0.1453
6	0.4955	0.8266	0.8792	0.8823	0.5968	0.5968
	0.2132	0.1565	<sup>b</sup> 0.1171	<sup>b</sup> 0.1179	0.1972	0.1995
7	0.4410	0.7638	0.7844	0.7849	0.6315	0.6307
	0.1470	0.1451	<sup>b</sup> 0.1445	<sup>b</sup> 0.1442	0.1910	0.1908
8	0.5719	0.8265	0.8484	0.8502	0.6852	0.6851
	0.1480	0.1362	<sup>b</sup> 0.1223	<sup>b</sup> 0.1223	0.1709	0.1705
9	0.4503	0.7826	0.8390	0.8428	0.4815	0.4820
	0.2013	0.1693	<sup>b</sup> 0.1446	<sup>b</sup> 0.1412	0.2339	0.2331
10	0.5411	0.9009	0.9025	0.9019	0.7834	0.7826
	0.1042	0.0692	0.0623	0.0623	0.1029	0.1026
11	0.6576	0.9356	0.9357	0.9358	0.8295	0.8297
	0.0992	0.0525	0.0521	0.0522	0.0957	0.0955
12	0.5656	0.9267	0.9431	0.9438	0.6095	0.6109
	0.1174	0.0680	<sup>b</sup> 0.0392	<sup>b</sup> 0.0365	0.1617	0.1607
13	0.5917	0.8271	0.8379	0.8416	0.7891	0.7886
	0.1229	0.0996	<sup>b</sup> 0.1018	<sup>b</sup> 0.1000	0.1061	0.1065
14	0.6923	0.8658	0.8738	0.8762	0.8254	0.8240
	0.1089	0.0839	<sup>b</sup> 0.0805	<sup>b</sup> 0.0802	0.0926	0.0930
15	0.6569	0.9330	0.9457	0.9473	0.7354	0.7354
	0.1451	0.0603	<sup>b</sup> 0.0475	<sup>b</sup> 0.0469	0.1140	0.1142
16	0.5596	0.8441	0.8601	0.8604	0.7153	0.7148
	0.1135	0.0984	<sup>b</sup> 0.0960	<sup>b</sup> 0.0957	0.1253	0.1259
17	0.6779	0.8874	0.8964	0.8970	0.7686	0.7678
	0.1082	0.0839	<sup>b</sup> 0.0805	<sup>b</sup> 0.0802	0.0926	0.0930
18	0.6237	0.9128	0.9309	0.9324	0.5858	0.5841
	0.1257	0.0645	<sup>b</sup> 0.0495	<sup>b</sup> 0.0486	0.1342	0.1355
19	0.6291	0.9439	0.9425	0.9424	0.8793	0.8795
	0.0791	0.0374	0.0373	0.0371	0.0618	0.0618
20	0.7540	0.9665	0.9648	0.9647	0.9183	0.9186
	0.0625	0.0254	0.0263	0.0265	0.0437	0.0436
21	0.6888	0.9688	0.9711 <sup>b</sup> 0.	0.9712	0.7824	0.7835
	0.0785	0.0244	0.0219	<sup>b</sup> 0.0216	0.0852	0.0843
22	0.6678	0.8576	0.8614	0.8632	0.8342	0.8325
	0.0867	0.0765	<sup>b</sup> 0.0765	<sup>b</sup> 0.0756	0.0784	0.0795
23	0.7689	0.8876	0.8896	0.8909	0.8672	0.8659
	0.0712	0.0549	<sup>b</sup> 0.0562	<sup>b</sup> 0.0560	0.0599	0.0603
24	0.7592	0.9672	0.9694	0.9699	0.8444	0.8439
	0.0937	0.0234	<sup>b</sup> 0.0211	<sup>b</sup> 0.0208	0.0611	0.0617
25	0.6509	0.8886	0.8958	0.8962	0.7956	0.7953
	0.0801	0.0576	<sup>b</sup> 0.0583	<sup>b</sup> 0.0581	0.0776	0.0778
26	0.7682	0.9282	0.9312	0.9314	0.8478	0.8472
	0.0708	0.0435	<sup>b</sup> 0.0433	<sup>b</sup> 0.0435	0.0658	0.0660
27	0.7368	0.9536	0.9574	0.9577	0.7192	0.7183
	0.0811	0.0345	<sup>b</sup> 0.0316	<sup>b</sup> 0.0315	0.0894	0.0890

<sup>a</sup>Values at the bottom of each row are standard deviations.

<sup>b</sup>Rejected  $H_0$  at  $\alpha = 0.000001$  level, where  $H_0$ : There is no difference between the mean correlation coefficient of simple cross-efficiency (SC-E) scores with the true efficiency scores and the mean correlation coefficient of efficiency scores of method-*i* with the true efficiency scores;  $H_a$ : The mean correlation coefficient of efficiency scores of method-*i* with the true efficiency scores is greater than the mean correlation coefficient of simple cross-efficiency (SC-E) scores with the true efficiency scores.

TABLE III: MEAN SPEARMAN'S RHO AND STANDARD DEVIATIONS<sup>a</sup> BETWEEN THE TRUE EFFICIENCY SCORES AND THE EFFICIENCY SCORES OF THE SIX METHODS

Case	Mean Spearman's rho: the true efficiency scores with efficiency scores of method- <i>i</i>					
	DEA	SC-E	BE-I	BE-II	AG-I	AG-II
1	0.4374 0.1878	0.7915 0.1439	0.7908 0.1359	0.7900 0.1382	0.6160 0.2005	0.6166 0.1988
2	0.5314 0.1665	0.8033 0.1189	0.8138 0.1171	0.8137 0.1163	0.6619 0.2032	0.6612 0.2028
3	0.4321 0.1863	0.8015 0.1450	0.82070. 1321	0.81930.1 341	0.4569 0.2515	0.4587 0.2505
4	0.5141 0.2044	0.7689 0.1524	0.78280. 1501	0.7889 <sup>b</sup> 0.1487	0.7125 0.1639	0.7124 0.1629
5	0.6091 0.1838	0.7785 0.1498	0.8001 <sup>b</sup> 0. 1335	0.8048 <sup>b</sup> 0.1320	0.7276 0.1505	0.7261 0.1511
6	0.5340 0.1785	0.7914 0.1479	0.8361 <sup>b</sup> 0.1218	0.8391 <sup>b</sup> 0.1208	0.5857 0.2086	0.5863 0.2084
7	0.4534 0.1899	0.7499 0.1515	0.7729 <sup>b</sup> 0.1491	0.7723 <sup>b</sup> 0.1490	0.6153 0.1950	0.6151 0.1942
8	0.5648 0.1686	0.7839 0.1343	0.8062 <sup>b</sup> 0.1183	0.8077 <sup>b</sup> 0.1194	0.6427 0.1825	0.6437 0.1814
9	0.4763 0.2013	0.7588 0.1689	0.8106 <sup>b</sup> 0.1461	0.8141 <sup>b</sup> 0.1395	0.4612 0.2501	0.4624 0.2476
10	0.5565 0.1322	0.8908 0.0689	0.8922 0.0619	0.8913 0.0625	0.7732 0.1125	0.7724 0.1124
11	0.6388 0.1149	0.8924 0.0588	0.8913 0.0571	0.8914 0.0565	0.7986 0.1030	0.7989 0.1027
12	0.5976 0.1186	0.8893 0.0576	0.9014 <sup>b</sup> 0.0302	0.9020 <sup>b</sup> 0.0274	0.6164 0.1676	0.6162 0.1676
13	0.6148 0.1229	0.8323 0.0996	0.8450 <sup>b</sup> 0.1018	0.8484 <sup>b</sup> 0.1000	0.7930 0.1061	0.7930 0.1065
14	0.6868 0.1224	0.8531 0.0758	0.8619 <sup>b</sup> 0.0719	0.8640 <sup>b</sup> 0.0715	0.8098 0.0870	0.8084 0.0876
15	0.6962 0.1143	0.8855 0.0595	0.8961 <sup>b</sup> 0.0462	0.8975 <sup>b</sup> 0.0455	0.7192 0.1242	0.7187 0.1242
16	0.5673 0.1344	0.8376 0.0976	0.8555 <sup>b</sup> 0.0930	0.8558 <sup>b</sup> 0.0925	0.7091 0.1307	0.7091 0.1311
17	0.6589 0.1215	0.8514 0.0741	0.8609 <sup>b</sup> 0.0730	0.8616 <sup>b</sup> 0.0728	0.7396 0.1084	0.7386 0.1088
18	0.6537 0.1152	0.8776 0.0578	0.8927 <sup>b</sup> 0.0409	0.8937 <sup>b</sup> 0.0408	0.5689 0.1462	0.5675 0.1472
19	0.6306 0.0936	0.9344 0.0373	0.9330 0.0383	0.9330 0.0382	0.8700 0.0650	0.8703 0.0645
20	0.7181 0.0696	0.9288 0.0352	0.9260 0.0356	0.9260 0.0356	0.8799 0.0432	0.8801 0.0434
21	0.7266 0.0760	0.9188 0.0145	0.9201 <sup>b</sup> 0.0137	0.9201 <sup>b</sup> 0. 0.122	0.7841 0.0861	0.7845 0.0851
22	0.6868 0.0867	0.8647 0.0765	0.8686 <sup>b</sup> 0.0765	0.8704 <sup>b</sup> 0.0756	0.8393 0.0784	0.8378 0.0795
23	0.7601 0.0760	0.8776 0.0477	0.8807 <sup>b</sup> 0.0490	0.8819 <sup>b</sup> 0.0489	0.8553 0.0539	0.8541 0.0541
24	0.7996 0.0682	0.9147 0.0199	0.9161 <sup>b</sup> 0.0185	0.9164 <sup>b</sup> 0.0184	0.8183 0.0633	0.8171 0.0641
25	0.6499 0.0801	0.8835 0.0576	0.8915 <sup>b</sup> 0.0583	0.8920 <sup>b</sup> 0.0581	0.7931 0.0776	0.7928 0.0778
26	0.7386 0.0755	0.8945 0.0448	0.8984 <sup>b</sup> 0.0445	0.8987 <sup>b</sup> 0.0448	0.8207 0.0623	0.8202 0.0625
27	0.7640 0.0735	0.9098 0.0249	0.9123 <sup>b</sup> 0.0227	0.9124 <sup>b</sup> 0.0225	0.7070 0.1018	0.7059 0.1015

<sup>a</sup>Values at the bottom of each row are standard deviations.

<sup>b</sup>Rejected  $H_0$  at  $\alpha=0.000001$  level, where  $H_0$ : There is no difference between the mean Spearman's rho of simple cross-efficiency (SC-E) scores with the true efficiency scores and the mean Spearman's rho of efficiency scores of method *i* with the true efficiency scores;  $H_a$ : The mean Spearman's rho of efficiency scores of method *i* with the true efficiency scores is greater than the mean correlation coefficient of simple cross-efficiency (SC-E) scores with the true efficiency scores.

In this simulation experiment, the results show that the three different weight sets,  $S_1$ ,  $S_2$  and  $S_3$  do not have any significant effect on the performance of the different methods. The differences in sample proportions of Advanced, Efficient, and Inefficient-DMUs also do not have significant effect on

the performance of the different methods.

## VI. CONCLUSION

This paper examines four models which apply secondary objectives to search for better weight sets in DEA. It attempts to examine the performance of those models in determining weight sets that are more accurate in the performance ranking of DMUs. Results of the simulation experiment show that BE-I and BE-II are the two best models in measuring performance of DMUs. Furthermore, the secondary objectives of both BE-I and BE-II do improve the accuracy of ranking DMUs in terms of their performance. The results also show that both aggressive models perform worse than simple cross-efficiency though their performance is better than DEA efficiency. These results can be explained by the following arguments. Secondary objectives search for better weight sets within the alternative optimal region. Some secondary objectives which are favor in choosing good weight sets like the benevolent models may choose weight sets that are more close to the underlying weight set while some secondary objectives like aggressive models which objectives are exactly opposite to the benevolent models may choose weight sets that are quite different from the underlying weight set. While simple cross-efficiency chooses weight set arbitrarily, hence its performance is in between benevolent models and aggressive models. Future research can explore more secondary objectives that can improve performance measure in DEA.

## REFERENCES

- [1] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of the decision making units," *European Journal of Operational Research*, vol. 2, pp. 429-444, 1978.
- [2] N. Adler, L. Friedman, and Z. Sinuany-Stern, "Review of ranking methods in the data envelopment analysis context," *European Journal of Operational Research*, vol. 140, pp. 249-265, 2002.
- [3] T. R. Sexton, R. H. Silkman, and A. J. Hogan, "Data envelopment analysis: critique and extensions," in *Measuring Efficiency: An Assessment of Data Envelopment Analysis*, R. H. Silkman, Eds., pp. 73-105, Jossey-Bass, San Francisco, 1986.
- [4] W. W. Cooper, J. L. Ruiz, and I. Sirvent, "Choosing weights from alternative optimal solutions of dual multiplier models in DEA," *European Journal of Operational Research*, vol. 180, pp. 443-458, 2007.
- [5] J. Doyle and R. Green, "Efficiency and cross-efficiency in DEA: derivations, meanings and uses," *Journal of the Operational Research Society*, vol. 45, pp. 567-578, 1994.
- [6] J. Doyle and R. Green, "Cross-evaluation in DEA: improving discrimination among DMUs," *INFOR*, vol. 33, pp. 205-222, 1995.
- [7] K. F. Lam, "In the determination of weight sets to compute cross-efficiency ratios in DEA," *Journal of the Operational Research Society*, vol. 61, pp. 134-143, 2010.
- [8] K. F. Lam and F. Bai, "Minimizing deviations of input and output weights from their means in data envelopment analysis," *Computers and Industrial Engineering*, vol. 60, pp. 527-533, 2011.



**Kim Fung Lam** received a BBA (1984), an MBA (1986) and a PhD (1992) in management science at Simon Fraser University, Vancouver, B.C., Canada.

He is an associate professor of Management Science at City University of Hong Kong, Hong Kong. He has published in journals such as *Annals of Operations Research*, *Computers and Industrial Engineering*, *Computers and Operations Research*, *European Journal of Operational Research*, *Journal of the Operational Research Society*, and *Personnel Review*, etc. His research interests include multi-criteria decision making, discriminant analysis, and data envelopment analysis. Dr. Lam is a member of the International Society on multiple criteria decision making.