RAIRO-Oper. Res. 53 (2019) 645–655
https://doi.org/10.1051/ro/2018060

RAIRO Operations Research
www.rairo-ro.org

GANGLESS CROSS-EVALUATION IN DEA: AN APPLICATION TO STOCK SELECTION

Gholam R. Amin[1,*] and Amar Oukil[2]

Abstract. This paper discusses the impact of ganging decision making units (DMUs) on the cross-efficiency evaluation in data envelopment analysis (DEA). A group of DMUs are said to be ganging-together if the minimum and the maximum cross-efficiency scores they give to all other DMUs are identical. This study demonstrates that the ganging phenomenon can significantly influence the cross-efficiency evaluation in favour of some DMUs. To overcome this shortcoming, we propose a gangless cross-efficiency evaluation approach. The suggested method reduces the effect of ganging and generates a more diversified list of top performing units. An application to the Tehran stock market is used to show the benefits of gangless cross-evaluation.

Mathematics Subject Classification. 90B30, 90B50, 90C05, 90C90.

Received April 26, 2017. Accepted July 7, 2018.

1. Introduction

Data envelopment analysis (DEA) is an optimization method for evaluating relative efficiencies of a group of homogenous decision making units (DMUs). Since its introduction in Charnes, Cooper and Rhodes [4], DEA has become one of the prominent areas in operations research for performance measurement and evaluation [8]. Recent years have seen a great variety of applications of DEA in almost every field (see Refs. [3, 12, 26]). The attractiveness of DEA lies in its ability to incorporate multiple input and multiple output factors without any underlying assumption of a functional form. The DEA self-evaluation has the flexibility to choose extremely high weights on some variables and extremely low weights on other variables in favour of the underlying DMU, which might hide the weaknesses of this unit. Meanwhile, the DEA cross-evaluation generates high scores for DMUs that are moderately good in all input and output variables [11]. Due to this desirable property, cross-efficiency evaluation is widely used in the DEA literature. Recent applications can be found in stock market portfolio selection [11], resource allocation [7], faculty academic performance evaluation [17], supply chain management [13], supplier selection under uncertainty [5], ranking of football players [20], and many others. In the meantime, the existence of alternate optimal solutions in DEA cross-evaluation motivated a lot of research on developing

Keywords. Data envelopment analysis, cross-efficiency evaluation, ganging decision making units, gangless cross-evaluation, stock selection.

[1] Faculty of Business, University of New Brunswick at Saint John, E2L 4L5 Saint John, NB, Canada.

[2] Department of Operations Management and Business Statistics, College of Economics and Political Science, Sultan Qaboos University, Muscat, Oman.

[*] Corresponding author: gamin@unb.ca

Article published by EDP Sciences

alternative secondary goal models (*e.g.*, [6,16]. Over the last decades, extensions of these models alongside new models have been developed (see, *e.g.*, [15,19,22,27] and [1,10,24]).

In the DEA cross-efficiency literature, the concept of ganging DMUs is initially addressed in [9,25]. According to [25], the ganging-together phenomenon occurs when DMUs with similar levels of inputs and outputs raise each other's scores due to the usage of similar weights over the peer-evaluation. Recently, [11] discussed the negative impact of ganging on diversification in portfolio selection and dealt with this issue through a mean-variance optimization model. Nonetheless, none of the aforementioned authors present a proper mathematical definition of ganging DMUs nor do they propose a sound approach to address the unwanted effects of this phenomenon. In this paper, we define more formally the concept of ganging DMUs within a peer-evaluation framework and we develop a procedure to identify potential gangs. Moreover, we show that a judicious removal of the duplicated scores of ganging DMUs from the evaluation process can contribute significantly to attenuate their concerted influence and, hence, generate a more diversified list of top performing cross-evaluation units. The proposed gangless cross-evaluation methodology is applied to select superior stocks in the Tehran stock market.

To the best of our knowledge, this paper is the first attempt to define formally the concept of ganging DMUs and propose an efficient algorithm to identify gangs. We illustrate the impact of ganging DMUs on the cross-efficiency evaluation through an example and propose a gangless cross-evaluation methodology to increase the fairness of the cross-evaluation. The proposed method is applied in the Tehran stock market to identify superior stocks.

The rest of this paper is organized as follows:

Section 2 illustrates the effect of ganging DMUs on the cross-evaluation process. In Section 3, we describe briefly the DEA cross-evaluation and we define the concept of ganging DMUs and we propose an efficient algorithm for gang identification. Section 4 presents an application of the proposed gangless cross-evaluation method to identify superior stocks in the Tehran stock market. Section 5 gives the concluding remarks.

2. Numerical illustration

To illustrate the impact of ganging DMUs on the cross-evaluation process, let us consider a sample of eight DMUs, each using a single input to produce two outputs, as shown in Table 1.

The corresponding production possibility set (PPS) is shown in Figure 1.

There are three efficient DMUs, A, B, and C, and five inefficient units, D–H. Each strong efficient DMU has two extreme optimal solutions which define an infinite set of alternative optimal solutions. These extreme optimal solutions correspond to the two line segments passing through each DMU. The following table presents the alternative optimal solutions for the three efficient DMUs.

As shown in Table 2, for instance, the optimal solutions for DMU-A are $(u_1^*, u_2^*) = (0, 0.3333)$ and $(\hat{u}_1^*, \hat{u}_2^*) = (0.0526, 0.3158)$, corresponding to the two line segments passing through A. The cross-efficiency scores of all DMUs, from the perspective of DMU-A, can be calculated using these two optimal solutions. Let us denote the cross-efficiency score of DMU-j from the perspective of DMU-k by e_{kj}, for any k and j. For example, e_{AB} can be calculated as follows

$$e_{AB}^* = (u_1^*, u_2^*)\mathbf{y}_B = (0, 0.3333) \times \begin{pmatrix} 2.5 \\ 2.75 \end{pmatrix} = 0.917$$

or

$$\hat{e}_{AB}^* = (\hat{u}_1^*, \hat{u}_2^*)\mathbf{y}_B = (0.0526, 0.3158) \times \begin{pmatrix} 2.5 \\ 2.75 \end{pmatrix} = 1$$

In DEA cross-evaluation, the non-uniqueness of cross-efficiency scores is a direct consequence of the existence of alternative optimal solutions in the DEA self-evaluation models. In fact, the cross-efficiency score of DMU-B, from the perspective of DMU-A, can be any value $e_{AB} \in [e_{AB}^{\min}, e_{AB}^{\max}]$ where $e_{AB}^{\min} = 0.917$ and $e_{AB}^{\max} = 1$. Tables 3A and 3B present the minimum and maximum cross-efficiency scores for all DMUs.

Column j $(j = A, \ldots, H)$ in the above two matrices, corresponds to the minimum and maximum cross-efficiency scores of DMU-j from the perspective of all other DMUs. Meanwhile, row i, $i = A, \ldots, H$, corresponds

TABLE 1. Eight DMUs with one input and two outputs.

DMUs	Input	Output 1	Output 2
A	1	1	3
B	1	2.5	2.75
C	1	3.5	1.5
D	1	3	0.5
E	1	2.75	0.85
F	1	2.5	0.5
G	1	2	0.7
H	1	1.85	0.6

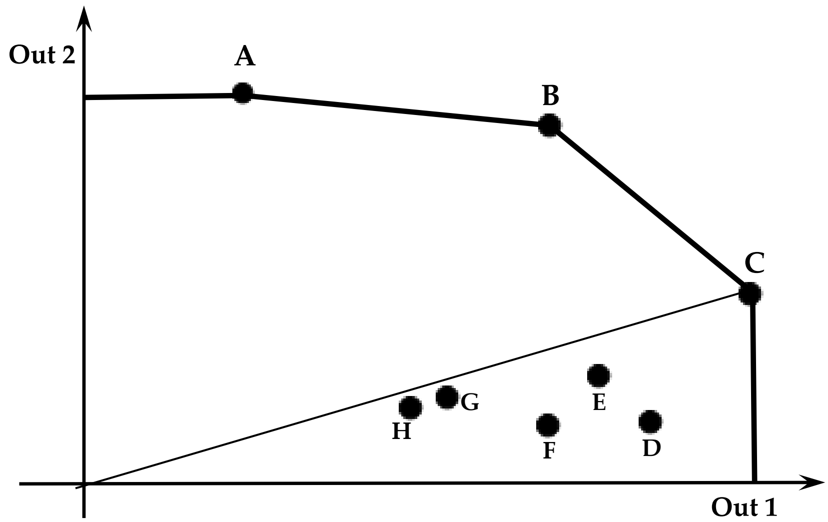

FIGURE 1. Illustration of ganging DMUs.

TABLE 2. Alternative optimal solutions.

DMUs	u_1^*	u_2^*	$\hat{u}_1^*$	$\hat{u}_2^*$
A	0.0000	0.3333	0.0526	0.3158
B	0.0526	0.3158	0.2128	0.1702
C	0.2128	0.1702	0.2857	0.0000

to the cross-efficiency scores of all DMUs from the perspective of DMU-i. The rows D–H are identical in both matrices because the self-evaluation DEA model corresponding to DMUs D–H has a unique optimal solution or, equivalently, there is a single line segment corresponding to these five DMUs. As a result, these DMUs give higher scores to all peers falling in their cluster [2, 21] and low scores to the others. This is one illustration of the ganging phenomenon [25] which benefits the DMUs that belong to the same cluster, such as DMU-C. On the opposite side, the real losers are undoutedly the DMUs that are far from this cluster, like efficienct DMU-A.

TABLE 3A. The Min cross-efficiency matrix.

	A	B	C	D	E	F	G	H
A	1.000	0.917	0.500	0.167	0.283	0.167	0.233	0.200
B	0.723	1.000	0.658	0.316	0.413	0.289	0.326	0.287
C	0.286	0.714	1.000	0.724	0.730	0.617	0.545	0.496
D	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
E	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
F	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
G	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
H	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529

TABLE 3B. The Max cross-efficiency matrix.

	A	B	C	D	E	F	G	H
A	1.000	1.000	0.658	0.316	0.413	0.289	0.326	0.287
B	1.000	1.000	1.000	0.724	0.730	0.617	0.545	0.496
C	0.723	1.000	1.000	0.857	0.786	0.714	0.571	0.529
D	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
E	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
F	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
G	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529
H	0.286	0.714	1.000	0.857	0.786	0.714	0.571	0.529

TABLE 4. The minimum and maximum (average) cross-efficiency scores.

DMUs	A	B	C	D	E	F	G	H
Min	0.430	0.775	0.895	0.687	0.670	0.580	0.495	0.454
Max	0.519	0.821	0.957	0.773	0.732	0.649	0.537	0.495

It is clear that the latter receives extremly low cross-efficieny scores from the ganging DMUs. Consequently, it would be unfair to consider these cross-efficieny scores as a ground for any cross-evaluation purpose. To illustrate this distortion, let us calculate the average cross-efficieny scores over the columns of Tables 3A and 3B.

As shown in Table 4, efficient DMU-A has the worst ranking position from both the pessimistic and the optimistic views. On the other hand, the inefficient DMUs that belong to the gangs have better positions than DMU-A. Moreover, DMU-C is ranked as the best and, hence, the winner in both cases.

To resolve this shortcoming, we propose eliminating $g-1$ ganging DMUs from the list of cross-efficieny evaluation before computing the averages, where g is the number of ganging DMUs. The removal of $g-1$ duplicated rows of scores can be viewed as keeping only a single representative voice for the whole gang. This leads to a gangless cross-evaluation with the results shown in Table 5.

Among the five ganging DMUs, D, E, F, G, and H, eliminating any subset of size four would generate the same cross-evaluation scores. As a result of the exclusion of the ganging DMUs, weak DMUs are disqualified from the top ranking positions. Furthermore, in the proposed gangless cross-evaluation, DMU-C is not the winner and DMU-A is not the loser anymore.

It is worth nothing that all inefficient DMUs are not necessarily in the class of ganging DMUs. It is also possible to observe some efficient units in the ganging DMUs. As shown in Figure 2, unlike the efficient

TABLE 5. Cross-efficiency (average) scores excluding gangs.

DMUs	A	B	C	D	E	F	G	H
Min	0.574	0.836	0.790	0.516	0.553	0.447	0.419	0.378
Max	0.752	0.929	0.915	0.689	0.679	0.584	0.503	0.460

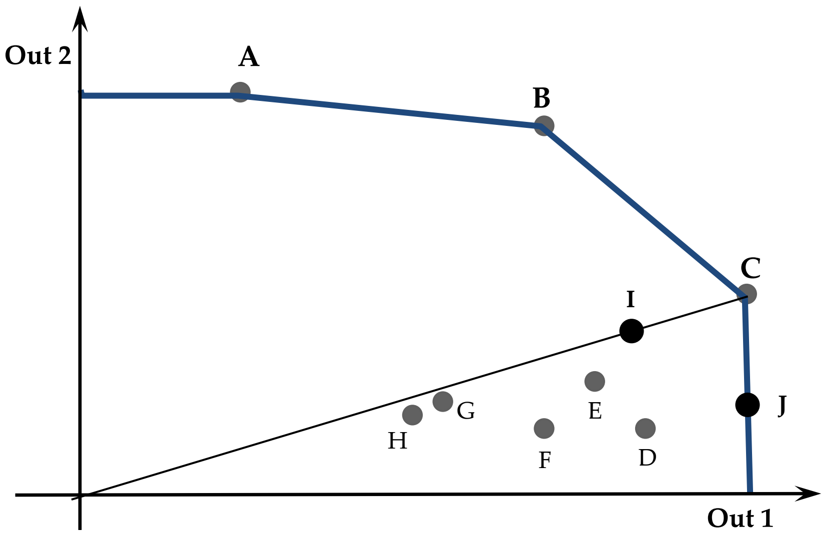

FIGURE 2. Non-ganging DMU-I and ganging DMU-J.

DMU-J which has a unique optimal solution and therefore a ganging unit with other inefficient DMUs, D–H, the inefficient DMU-I has two different extreme optimal solutions and hence a non-ganging unit.

It is also worthwhile noting that gangs' elimination allows also a more diversified list of top performing DMUs.

3. Gangless cross-evaluation

Suppose there are n DMUs, each using m inputs and producing s outputs. Let us also assume $\mathbf{x}_j = (x_{1j},\ldots,x_{mj})$ and $\mathbf{y}_j = (y_{1j},\ldots,y_{sj})$ are, respectively, the vectors of inputs and outputs for DMU-j (j=1,...,n). The self-evaluation DEA model for DMU-k can be written as follows [4]:

$$\begin{aligned} & e^*_{kk} = \max \sum_{r=1}^{s} y_{rk}u_r \\ & \text{s.t.} \\ & \sum_{i=1}^{m} x_{ik}v_i = 1 \\ & \sum_{r=1}^{s} y_{rj}u_r - \sum_{i=1}^{m} x_{ij}v_i \le 0 j = 1,\ldots,n \\ & u_r \ge 0 r = 1,\ldots,s, v_i \ge 0 i = 1,\ldots,m \end{aligned} \tag{3.1}$$

where $\mathbf{v} = (v_1,\ldots,v_m)$ and $\mathbf{u} = (u_1,\ldots,u_s)$ are, respectively, decision variables representing the weights of inputs and outputs. The efficiency score of DMU-k,e^*_{kk}, as well as the cross-efficiency scores of the other DMUs from the perspective of DMU-k can be obtained from optimal solution(s) of model (3.1). More precisely, assuming $(\mathbf{u}^*,\mathbf{v}^*) = (u^*_1,\ldots,u^*_s,v^*_1,\ldots,v^*_m)$ as an optimal solution of model (3.1), the efficiency score of DMU-k

is $e^*_{kk} = \mathbf{u}^*\mathbf{y}_k = \sum_{k=1}^{s} u^*_r y_{rk}$ and the cross-efficiency score of DMU-l from the perspective of DMU-k is obtained as follows:

$$e_{kl} = \frac{\mathbf{u}^*\mathbf{y}_l}{\mathbf{v}^*\mathbf{x}_l}, \quad \forall l = 1, \ldots, n, l \neq K$$

A cross-efficiency matrix $\mathbf{E} = (e_{kl})_{k,l=1,\ldots,n}$ is generated by calculating the cross-efficiency scores of all DMUs from the perspective of all other units. In this matrix, column $j \quad (j = 1, \ldots, n)$ represents the cross-efficiency scores of DMU-j from the perspective of all other DMUs. The average of this column, denoted by $\bar{e}_j$, is simply defined as the cross-efficiency score of DMU-j $(j = 1, \ldots, n)$. Moreover, row $i \quad (i = 1, \ldots, n)$ in matrix $\mathbf{E}$ represents the cross-efficiency scores of all DMUs from the perspective of DMU-i. The matrix of cross-efficiency scores $\mathbf{E}$ is not unique and alternative optimal solutions existing in the self-evaluation DEA model generate different cross-efficiency matrices. Assume $\mathbf{E}^{\min} = \left(e^{\min}_{kl}\right)_{n\times n}$ and $\mathbf{E}^{\max} = \left(e^{\max}_{kl}\right)_{n\times n}$ denote, respectively, the matrices of minimum and maximum cross-efficiency, where the minimum and maximum cross-efficiency scores of DMU-l from the perspective of DMU-k can be obtained by solving the following DEA models [15, 19].

$$\begin{aligned}
& e^{\min}_{kl} = \min \sum_{r=1}^{s} y_{rl} u_r \\
& \text{s.t.} \\
& \sum_{i=1}^{m} x_{il} v_i = 1 \\
& \sum_{r=1}^{s} y_{rj} u_r - \sum_{i=1}^{m} x_{ij} v_i \leq 0 j = 1, \ldots, n, j \neq K \\
& \sum_{r=1}^{s} y_{rk} u_r - e^*_{kk} \sum_{i=1}^{m} x_{ik} v_i = 0 \\
& u_r \geq 0 r = 1, \ldots, s, v_i \geq 0 i = 1, \ldots, m
\end{aligned} \tag{3.2}$$

$$\begin{aligned}
& e^{\max}_{kl} = \max \sum_{r=1}^{s} y_{rl} u_r \\
& \text{s.t.} \\
& \sum_{i=1}^{m} x_{il} v_i = 1 \\
& \sum_{r=1}^{s} y_{rj} u_r - \sum_{i=1}^{m} x_{ij} v_i \leq 0 j = 1, \ldots, n, j \neq K \\
& \sum_{r=1}^{s} y_{rk} u_r - e^*_{kk} \sum_{i=1}^{m} x_{ik} v_i = 0 \\
& u_r \geq 0 r = 1, \ldots, s, v_i \geq 0 i = 1, \ldots, m
\end{aligned} \tag{3.3}$$

One of the shortcomings in the use of DEA cross-evaluation is the existence of ganging DMUs. This concept is introduced in [25] without presenting a proper mathematical definition nor a sound approach to address its unwanted effects on the cross-evaluation methods. Based on the discussion in Section 2, we propose the following definition.

Definition (**Ganging DMUs**): In cross-efficiency evaluation, a group of (at least two) DMUs, indexed in G, are said to be ganging-together if and only if $e^{\min}_{pl} = e^{\min}_{ql} = e^{\max}_{pl} = e^{\max}_{ql}$, for each $p, q \in G$ and $l = 1, \ldots, n$.

In the following we propose a two-stage procedure to identify ganging DMUs.

Stage 1: Identify potential gangs

Let $r^{\min}_i$ and $r^{\max}_i$ denote the sum of cross-efficiency scores of the ith row in matrices $\mathbf{E}^{\min}$ and $\mathbf{E}^{\max}$, respectively, *i.e.*, $r^{\min}_i = \sum_{l=1}^{n} e^{\min}_{il}$ and $r^{\max}_i = \sum_{l=1}^{n} e^{\max}_{il}$.

A necessary condition for DMUs p and q to be ganging-together is

$$r^{\min}_p = r^{\max}_p = r^{\min}_q = r^{\max}_q \tag{3.4}$$

Accordingly, a set G is said to be a potential gang if condition (3.4) holds for any pair of DMUs in G. Therefore, $r_i^{\min} = r_i^{\max} = r$ for all $i \in G$. Without loss of generality, let us assume that $G = \{1, 2, \ldots, g\}$ denotes a set of potential ganging DMUs. It is not difficult to see that the number of such potential gangs G is bounded by the number of clusters [2, 21] in the corresponding PPS.

In Stage 2, any potential gang G will be checked for satisfying the sufficient condition given in the definition of ganging DMUs. This may produce more than one gang.

Stage 2: Gang's refinement

Consider the sub-matrix M_G of cross-efficiency scores corresponding to G in $\mathbf{E}^{\min}$, and assume that $R_k^{\min}$ (for all $k \in G$) denotes the rows of M_G.

With the above notations, we propose the following procedure to identify the ganging DMUs from G.

Let *Count* be the counter of rows $R_k^{\min}$ ($k \in$ List) that are similar to $R_1^{\min}$, *i.e.*, the rows that satisfy $R_k^{\min} - R_1^{\min} = \underline{0}$, where $\underline{0}$ is the row whose entries are all zero, and *List* is the set of rows of M_G that are yet to be assessed at a given iteration.

Algorithm:
Input: A sub-matrix $M_G = (R_k^{\min})_{k \in G}$ of cross-efficiency scores.
Output: A set of gangs.
begin
do {

Step 0: Initialization
- $List = \text{G}$
- Assign a DMU identifier $ID(i)$ to each row index i, $i = 1, \ldots, g$
- $Count = 0$

Step 1: Construct the similarity matrix SM

$$SM = \begin{pmatrix} R_1^{\min} - R_1^{\min} \\ R_2^{\min} - R_1^{\min} \\ \vdots \\ R_g^{\min} - R_1^{\min} \end{pmatrix}$$

Step 2: Compute *Count*, the number of rows in SM that satisfy $R_k^{\min} - R_1^{\min} = \underline{0}$

Step 3: Assess the value of *Count*

if $Count = g$ **then**
- All DMUs in *List* are ganging-together.
- **STOP.**

if $Count = 1$ **then**
- DMU $ID(1)$ is an *intruder* that must be discarded.
- Update *List*: $List \leftarrow List \setminus \{1\}$
- Decrement g: $g \leftarrow g - 1$
- **goto** Step 0.

if $Count = z$ and $2 \leq z \leq g$ **then**
- A group C_z of z ganging DMUs is identified.
- Update *List*: $List \leftarrow List \setminus C_z$
- Decrement g: $g \leftarrow g - z$
- **goto** Step 0. **}**

while ($List \neq \phi$)
Output all the groups of ganging DMUs.
end

Note that the idea of Assigning a DMU identifier $ID(i)$ to each row index i in Step 0 consists in running the algorithm at each iteration with rows i from 1 to g while keeping in memory the DMU that is associated to each row i.

The following theorem shows that the proposed gang identification algorithm is polynomial.

Theorem 3.1. *The complexity of the algorithm is $O(n^4)$.*

Proof. The complexity of the algorithm is equal to the complexity of Stage 2. This stage needs to be run at most C times, where C is the number of clusters in the PPS that can be obtained by the algorithm suggested in [21]. Clearly, $\mathrm{C} \leq n+1$. In addition, each potential gang may consist of at most $n-1$ DMUs, *i.e.*, $g \leq n-1$. Therefore, the total number of operations required to run the proposed algorithm is $O(n^4)$. This completes the proof. □

The implementation of the gangless cross-evaluation is based on the elimination of the ganging DMUs identified through the above algorithm.

4. An application to the Tehran Stock Market

This section demonstrates the impact of ganging DMUs on the selection of top units. We use a dataset from the Tehran Stock Market (TSM) consisting of 23 manufacturing stocks. According to the raw financial data that are available in the annual financial statement, we consider two inputs and two outputs. Following the research studies in the literature [11, 14, 23], the two inputs are Leverage (LEV) and Cash ratios and the two outputs are Return on assets (ROA) and Earnings per share (EPS). Table 6 gives the dataset for 23 stocks with their self-evaluation DEA scores.

For illustration purposes, the dataset in Table 6 is taken from the financial year 2011, even if the proposed methodology can be easily extended to the analysis of manufacturing stocks covering several years. In portfolio selection, it is important to invest in a diversified cluster of stocks in order to reduce the risk of losing the whole investment.

Using models (3.2) and (3.3), we generate the cross-efficiency matrices $\mathbf{E}^{\min} = \left(e_{kl}^{\min}\right)_{n\times n}$ and $\mathbf{E}^{\max} = \left(e_{kl}^{\max}\right)_{n\times n}$ corresponding to the 23 stocks in the TSM. Using the algorithm proposed in this paper, there are two ganging groups for this application:

$$\mathrm{G}_1 = \{1, 3, 7, 8, 11, 20, 21\}$$
$$\mathrm{G}_2 = \{2, 13, 16, 18\}$$

The following table shows the top 5 selected stocks in two cases: in the presence of the two ganging groups and in the absence of the ganging DMUs.

As shown in Table 7, stock company S_{10} is in the third top position. This top position is obtained with the support of the seven ganging DMUs indexed in G_1. On the other hand, S_{10} is not in the top five stocks when the proposed gangless cross-evaluation is used. A similar scenario is reproduced for stock company S_{11}. In the presence of its supporting ganging DMUs in G_1, S_{11} ranks fifth however, it is not the top of the gangless cross-evaluation. This shows the impact of ganging DMUs exiting in G_1, as S_{10} and S_{11} are in the same cluster [2, 21] of ganging DMUs in G_1 and uses their support.

On the other hand, the stock companies S_{23} and S_{19} are not in the top five in the presence of ganging DMUs. These two stocks have the opportunity to be in the top five stocks when ganging DMUs are eliminated using the gangless cross-efficiency method.

Although S_{15} conserves its top rank position within both frameworks, it is worthwhile noting that its ultimate cross-evaluation score decreases after losing the support of its gang in G_2. On the contrary, S_{12} which belongs to neither gang and, hence, was previously penalized, gains more value after discarding the ganging DMUs in the gangless cross-evaluation procedure.

TABLE 6. Data for 23 stocks in the Tehran stock market.

	Inputs		Outputs		
DMUs	Cash	LEV	EPS	ROA	DEA Scores
S_1	0.2485	0.4688	4.00	0.2002	0.3712
S_2	0.1284	0.6539	3.70	0.0902	0.2893
S_3	0.293	0.5644	4.20	0.1627	0.2531
S_4	0.0677	0.5697	3.00	0.0903	0.4826
S_5	0.0446	0.7492	0.35	0.0083	0.0631
S_6	0.0353	0.5411	0.85	0.0494	0.4400
S_7	0.2886	0.4385	6.20	0.2578	0.4564
S_8	0.3103	0.567	6.00	0.2126	0.3207
S_9	0.0295	0.8002	1.26	0.0524	0.4999
S_{10}	0.2936	0.1464	10.00	0.4006	1.0000
S_{11}	0.2871	0.211	9.80	0.3925	0.9104
S_{12}	0.1011	0.4994	10.00	0.2614	1.0000
S_{13}	0.0678	0.8587	2.00	0.0256	0.2216
S_{14}	0.0429	0.4101	2.50	0.1274	1.0000
S_{15}	0.0812	0.192	7.00	0.1988	1.0000
S_{16}	0.1067	0.8744	2.30	0.0424	0.1902
S_{17}	0.1165	0.7843	0.60	0.0226	0.0704
S_{18}	0.0976	0.6923	4.00	0.0837	0.3779
S_{19}	0.0216	0.5669	1.65	0.0773	1.0000
S_{20}	0.2164	0.4075	8.00	0.2394	0.5101
S_{21}	0.2451	0.4021	5.50	0.2041	0.4107
S_{22}	0.2408	0.3492	10.60	0.2983	0.7258
S_{23}	0.0257	0.885	5.90	0.0808	1.0000

TABLE 7. Stock selection: Gangs *vs.* Gangless.

	Cross-evaluation with gangs		Gangless cross-evaluation	
Rank	Stocks	Max-Cross	Stocks	Max-Cross
1	S_{15}	0.9759	S_{15}	0.966
2	S_{12}	0.8491	S_{12}	0.8834
3	S_{10}	0.7163	S_{14}	0.7903
4	S_{14}	0.6765	S_{23}	0.7269
5	S_{11}	0.673	S_{19}	0.7206

Henceforth, dropping ganging DMUs increases fairness of the assessment process and reduces misjudgment. Nonetheless, the effect of such a decision may be attenuated if the ganging DMUs are uniformly distributed across the different clusters of the PPS, as shown in the figure.

Figure 3 presents the PPS corresponding to dataset 1 taken from [11]. In this illustrative example, [11] considered twenty two DMUs each with two inputs and a single output. Using the gang identification algorithm proposed in this paper, we extract the following three gangs.

$$\mathrm{G}_1 = \{1, 2, 4, 16, 17, 19\}, \mathrm{G}_2 = \{3, 5, 7, 10, 13, 15\}, G_3 = \{8, 9, 12\}$$

As shown in Figure 3, there are six DMUs in the first two gangs and three DMUs in the last one. With an equal number of ganging DMUs, the gangless cross-efficiency evaluation will definitely have less effect on the cross-efficiency scores.

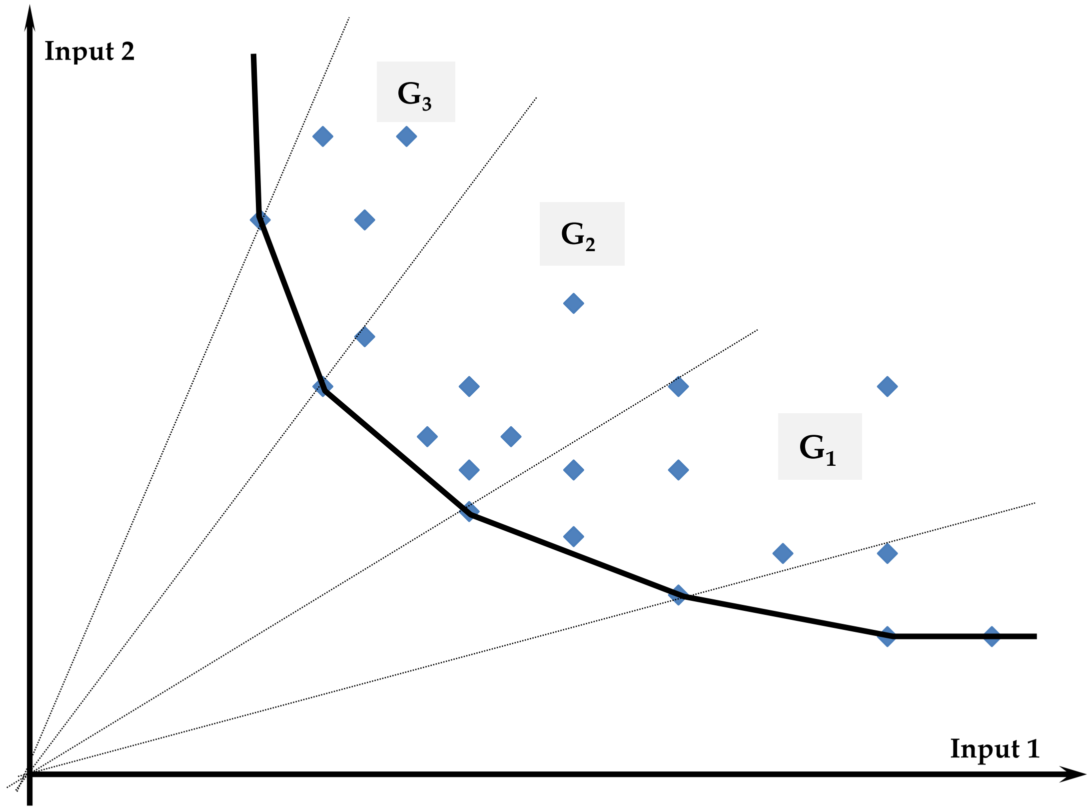

FIGURE 3. Three ganging groups.

5. Concluding remarks

The impact of ganging decision making units (DMUs) on the cross-evaluation in data envelopment analysis is investigated in this paper. A formal definition of the ganging concept is provided and an efficient algorithm is developed to identify the ganging groups. Furthermore, a gangless cross-evaluation methodology is suggested to reduce the effect of the ganging phenomenon. It is shown that, by eliminating the ganging DMUs, the proposed gangless cross-evaluation generates more diversified lists of top performing units. An application to the Tehran Stock Market is used to show the applicability of the proposed method.

The proven importance of the ganging concept in decision making may stimulate research on investigating the phenomenon from the weight profiles perspective under a cross-evaluation framework [18]. Such a prospective direction could be interesting as it may look at the interaction among the ganging DMUs at the source (weight profiles) rather than restricting the analysis within the cross-efficiency scores. In addition, further research is required to investigate the impact of gangs on the cross-evaluation by incorporating efficient and inefficient units in the definition of the ganging DMUs.

References

[1] J. Alcaraz, N. Ramón, J.L. Ruiz and I. Sirvent, Ranking ranges in cross-efficiency evaluations. *Eur. J. Oper. Res.* **226** (2013) 516–521.

[2] G.R. Amin, A. Emrouznejad and S. Rezaei, Some clarifications on the DEA clustering approach. *Eur. J. Oper. Res.* **215** (2011) 498–501.

[3] M. Asmild, J.C. Paradi and J.T. Pastor, DEA based models for reallocations of police personnel. *OR Spect.* **34** (2012) 921–941.

[4] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[5] M. Dotoli, N. Epicoco, M. Falagario and F. Sciancalepore, A stochastic cross-efficiency data envelopment analysis approach for supplier selection under uncertainty. *Int. Trans. Oper. Res.* **23** (2016) 725–748.

[6] J. Doyle and R. Green, Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *J. Oper. Res. Soc.* **45** (1994) 567–578.

[7] J. Du, W.D. Cook, L. Liang and J. Zhu, Fixed cost and resource allocation based on DEA cross-efficiency. *Eur. J. Oper. Res.* **235** (2014) 206–214.

[8] A. Emrouznejad, Advances in data envelopment analysis. *Ann. Oper. Res.* **214** (2014) 1–4.

[9] R. Green and J. Doyle, Improving discernment in DEA using profiling: a comment. *Omega* **24** (1996) 365–366.

[10] B.H. Jeong and C.S. Ok, A new ranking approach with a modified cross efficiency matrix. *Asia-Pac. J. Oper. Res.* **30** (2013) 1–17.

[11] S. Lim, K.W. Oh, and J. Zhu, Use of DEA cross-efficiency evaluation in portfolio selection: an application to Korean stock market. *Eur. J. Oper. Res.* **236** (2014) 361–368.

[12] J.S. Liu, L.Y. Lu and W.M. Lu, Research fronts in data envelopment analysis. *Omega* **58** (2016) 33–45.

[13] M. Mahdiloo, R.F. Saen and K.H. Lee, Technical, environmental and eco-efficiency measurement for supplier selection: an extension and application of data envelopment analysis. *Int. J. Prod. Econ.* **168** (2015) 279–289.

[14] Z. Mashayekhi and H. Omrani, An integrated multi-objective Markowitz-DEA cross-efficiency method with fuzzy returns for portfolio selection problem. *Appl. Soft Comput.* **38** (2016) 1–9.

[15] M. Oral, G.R. Amin and A. Oukil, Cross-efficiency in DEA: a maximum resonated appreciative model. *Measurement* **63** (2015) 159–167.

[16] M. Oral, O. Kettani and P. Lang, A methodology for collective evaluation and selection of industrial R&D projects. *Manage. Sci.* **37** (1991) 871–885.

[17] M. Oral, A. Oukil, J.-L. Malouin and O. Kettani, The appreciative democratic voice of DEA: a case of faculty academic performance evaluation. *Socio-Econ. Planning Sci.* **48** (2014) 20–28.

[18] A. Oukil, Ranking via composite weighting schemes under a DEA cross-evaluation framework. *Comput. Ind. Eng.* **117** (2018) 217–224.

[19] A. Oukil and G.R. Amin, Maximum appreciative cross-efficiency in DEA: a new ranking method. *Comput. Ind. Eng.* **81** (2015) 14–21.

[20] A. Oukil and S.M. Govindaluri, A systematic approach for ranking football players within an integrated DEA-OWA framework. *Manage. Decis. Econ.* **38** (2017) 1125–1136.

[21] R.W. Po, Y.Y. Guh and M.S. Yang, A new clustering approach using data envelopment analysis. *Eur. J. Oper. Res.* **199** (2009) 276–284.

[22] N. Ramón, J.L. Ruiz and I. Sirvent, Dominance relations and ranking of units by using interval number ordering with cross-efficiency intervals. *J. Oper. Res. Soc.* **65** (2014) 1336–1343.

[23] K.Y. Shen and G.H. Tzeng, Combined soft computing model for value stock selection based on fundamental analysis. *Appl. Soft Comput.* **37** (2015) 142–155.

[24] M. Soltanifar and S. Shahghobadi, Selecting a benevolent secondary goal model in data envelopment analysis cross-evaluation by a voting model. *Socio-Econ. Planning Sci.* **47** (2013) 65–74.

[25] C. Tofallis, Improving discernment in DEA using profiling. *Omega* **24** (1996) 361–364.

[26] Y.M. Wang, Y. Luo and Y.S. Xu, Cross-weight evaluation for pairwise comparison matrices. *Group Decis. Negotiat.* **22** (2013) 483–497.

[27] S. Washio and S. Yamada, Evaluation method based on ranking in data envelopment analysis. *Expert Syst. Appl.* **40** (2013) 257–262.