Preference aggregation and DEA: An analysis of the methods proposed to discriminate efficient candidates

Bonifacio Llamazares *, Teresa Peña

Dep. de Economía Aplicada, PRESAD Research Group, Universidad de Valladolid, Avda. Valle de Esgueva 6, 47011 Valladolid, Spain.

Abstract

There are different ways to allow the voters to express their preferences on a set of candidates. In ranked voting systems, each voter selects a subset of the candidates and ranks them in order of preference. A well-known class of this voting systems are scoring rules, where fixed scores are assigned to the different ranks and the candidates with the highest score are the winners. One of the most important issues in this context is the choice of the scoring vector, since the winning candidate can vary according to the scores used. To avoid this problem, Cook and Kress [Management Science 36, pp. 1302-1310, 1990], using a DEA/AR model, proposed to assess each candidate with the most favorable scoring vector for him/her. However, the use of this procedure often causes several candidates to be efficient, i.e., they achieve the maximum score. For this reason, several methods to discriminate among efficient candidates have been proposed. The aim of this paper is to analyze and show some drawbacks of these methods.

Key words: Scoring rules, Data envelopment analysis, Discrimination of efficient candidates.

1 Introduction

An important issue in the decision-making framework is how to obtain a social ranking or a winning candidate from individuals' preferences on a set of candidates $\{A_1, \ldots, A_m\}$. In some voting systems, each voter selects k

* Corresponding author.
Email addresses: boni@eco.uva.es (Bonifacio Llamazares), maitepe@eco.uva.es (Teresa Peña).

candidates and ranks them from most to least preferred (*ranked voting systems*). Among these systems, a well-known procedure to obtain a social ranking or a winning candidate are *scoring rules*, where fixed scores are assigned to the different ranks. In this way, the score obtained by the candidate A_i is $Z_i = \sum_{j=1}^{k} w_j v_{ij}$, where v_{ij} is the number of j-th place ranks that candidate A_i occupies and $(w_1, \ldots, w_k)$ is the scoring vector used. The plurality rule, where $w_1 = 1$ and $w_j = 0$ for all $j \in \{2, \ldots, k\}$, and the Borda rule, where $k = m$ and $w_j = m - j$ for all $j \in \{1, \ldots, m\}$, are the best known examples of scoring rules.

It is worth noting that the Borda rule has interesting properties in relation to other scoring rules. According to Brams and Fishburn [2, p. 226]:

> "Among ranked positional scoring procedures to elect one candidate, Borda's method is superior in many respects, including susceptibility to strategic manipulation, propensity to elect Condorcet candidates, and ability to minimize paradoxical possibilities".

Nevertheless, there are numerous decisional contexts where the scoring rules used are different from the Borda rule. In these cases, the choice of the scoring vector used is somewhat arbitrary.

On the other hand and irrespective of the scoring vector used, the utilization of a fixed scoring vector presents the following drawback: a candidate that is not the winner with the scoring vector imposed initially could be so if another scoring vector is used. To avoid this problem, Cook and Kress [3] suggested evaluating each candidate with the most favorable scoring vector for him/her. With this purpose, they introduced Data Envelopment Analysis (DEA) in this context. The model DEA/AR proposed by these authors is:

$$
\begin{aligned}
Z_o^* = \max\ & \sum_{j=1}^{k} w_j v_{oj}, \\
\text{s.t.}\ & \sum_{j=1}^{k} w_j v_{ij} \le 1, \quad i = 1, \ldots, m, \\
& w_j - w_{j+1} \ge d(j, \varepsilon), \quad j = 1, \ldots, k-1, \\
& w_k \ge d(k, \varepsilon),
\end{aligned}
\tag{1}
$$

where $\varepsilon \ge 0$ and the functions $d(j, \varepsilon)$, called the *discrimination intensity functions*, are nonnegative and nondecreasing in ε. Furthermore, $d(j, 0) = 0$ for all $j \in \{1, \ldots, k\}$.

The principal drawback of this procedure is that several candidates are often efficient, i.e., they achieve the maximum attainable score ($Z_o^* = 1$). To avoid this weakness, Cook and Kress [3] proposed to maximize the gap between consecutive weights of the scoring vector and, in this way, to reduce the feasible set of problem (1). Thus, the model considered by these authors is:

$$
\begin{aligned}
&\max\ \varepsilon, \\
&\text{s.t.}\ \sum_{j=1}^{k} w_j v_{ij} \leq 1, \quad i = 1, \ldots, m, && \text{(a)} \\
&w_j - w_{j+1} \geq d(j, \varepsilon), \quad j = 1, \ldots, k-1, && \\
&w_k \geq d(k, \varepsilon), && \text{(b)}
\end{aligned}
\tag{2}
$$

where ε and the functions $d(j, \varepsilon)$ satisfy the conditions imposed in Model (1). Cook and Kress [3] demonstrated that, at optimality, at least one of the constraints in (2a) and all the constraints in (2b) hold as equalities. The candidate(s) A_i for which $\sum_{j=1}^{k} w_j v_{ij} = 1$ are the winning candidates (see Cook and Kress [3, p. 1308]). The candidate(s) in second place can be found by re-solving Model (2) after deleting the binding constraint(s) from (2a). This process can be repeated until the order of all candidates is fixed.

However, Green *et al.* [5] noticed two important drawbacks of the previous procedure. The first one is that the choice of the functions $d(j, \varepsilon)$ is not obvious, and that choice determines the winner. The second one is that for an important class of discrimination intensity functions (that for which $d(j, \varepsilon) = g(j)h(\varepsilon)$, with $h(\varepsilon)$ strictly monotonic increasing) the previous procedure amounts to the scoring rule given by $w_j = \sum_{u=j}^{k} g(u)$. Therefore, when Cook and Kress's model is used with this class of discrimination intensity functions, the aim pursued by these authors (evaluating each candidate with the most favorable scoring vector for him/her) is not reached.

Due to the drawbacks mentioned above, other procedures to discriminate efficient candidates have appeared in the literature. The aim of this paper is to show some problems that these procedures present. In Section 2 we analyze the model proposed by Green *et al.* [5], who suggest a discrimination method using a cross-evaluation matrix instead. In Section 3 we study the model given by Hashimoto [6], which uses super-efficiency technique and assumes the condition of decreasing and convex sequence of weights. Section 4 is devoted to the model proposed by Noguchi *et al.* [7]. These authors use decreasing sequences of weights in the model of Green *et al.* [5]. In Section 5 we analyze the model suggested by Obata and Ishii [8]. They consider that, in order to compare the maximum score obtained by each candidate, it is fair to use scoring vectors of the same size. So, they suggest normalizing the most favorable scoring vectors for each efficient candidate. For this model, we give a characterization to determine the winning candidates when the L_∞-norm or the L_1-norm are used. This characterization allows us to obtain the winning candidates without solving the programming model proposed by Obata and Ishii [8]. Finally, Section 5 summarizes the main problems of the methods analyzed in this paper.

2 Green, Doyle and Cook's model

To avoid the problem of the choice of discrimination intensity functions, Green *et al.* [5] consider $d(j, \varepsilon) = 0$ for all $j \in \{1, \ldots, k\}$. Moreover, these authors propose to use the cross-evaluation method, introduced by Sexton *et al.* [9], to discriminate efficient candidates. In the cross-evaluation method the Model (1) is solved for each candidate A_i. Next, the values $h_{iq} = \sum_{j=1}^{k} w_j v_{qj}$, $q = 1, \ldots, m$, which represents the score obtained by candidate A_q when evaluated with some of the most favorable scoring vectors for candidate A_i, are calculated. Given that the values h_{iq} depend on the scoring vector used, Green *et al.* [5], according to Sexton *et al.* [9], propose two ways to select this vector: the benevolent cross-evaluation and the aggressive cross-evaluation. In the benevolent context, the scoring vector that maximizes the value h_{iq} is selected. However, in the aggressive context, the scoring vector that minimizes the value h_{iq} is chosen. It is worth noting that this version provides greater discrimination between the candidates than the benevolent. Furthermore, the aggressive version seems to be more logical from the point of view of every candidate, who will want his/her adversaries to obtain the smallest possible score.

Once calculated, the values h_{iq} are summarized in a matrix H of order $m \times m$. The values of row i represent the score obtained by each candidate when evaluated by A_i, and the values of column q represent the score obtained by A_q when evaluated by all candidates. In order to obtain a global score for each candidate, Green *et al.* [5] initially propose, according to Sexton *et al.* [9], the average of the scores obtained by each candidate; i.e., $h_q = (1/m) \sum_{i=1}^{m} h_{iq}$ is the global score given to candidate A_q. Nevertheless, in the preference aggregation framework this procedure presents a serious drawback: the number of first, second, ..., k-th ranks obtained by inefficient candidates may change the order of efficient candidates. To mitigate this problem, Green *et al.* [5] suggest that, in the establishment of the global scores, the weight of each candidate be proportional to his/her global score instead of $1/m$. Specifically, if $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_m)$ is the row vector of global scores, these authors propose to solve the equation:

$$\left(1/\sum_{i=1}^{m}\Theta_i\right)\boldsymbol{\Theta}H=\boldsymbol{\Theta}, \tag{3}$$

to obtain Θ_i, candidate A_i's global score. For this, the following iterative procedure can be used:

Step 1: $\boldsymbol{\Theta}^{new}=\left(1/\sum_{i=1}^{m}\Theta_i{}^{old}\right)\boldsymbol{\Theta}^{old}H,$

Step 2: $\boldsymbol{\Theta}^{old}=\boldsymbol{\Theta}^{new},$

where initially $\boldsymbol{\Theta}^{old}=(1,\ldots,1)$.

Nevertheless, the problem described above still remains when we use this method to obtaining the global scores. This situation can be illustrated with the following example.

Example 1. Table 1 shows the number of first and second ranks obtained by three candidates.

Table 1
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	7	1
B	4	7
C	3	6

When Model (1) is solved for each of the three candidates, the following scores are obtained:

$$Z^*_{\mathrm{A}}=1,\qquad Z^*_{\mathrm{B}}=1,\qquad Z^*_{\mathrm{C}}=9/11.$$

Therefore, candidates A and B are efficient. Table 2 shows the aggressive cross-evaluation matrix corresponding to the Table 1 data.

The candidates' global scores are the following:

$$\Theta_{\mathrm{A}}=0.823,\qquad \Theta_{\mathrm{B}}=0.850,\qquad \Theta_{\mathrm{C}}=0.675.$$

So, candidate B is the winner. Let us assume that candidate C loses three first ranks to a new candidate D without any variation in the remaining can-

Table 2
Aggressive cross-evaluation matrix for the Table 1 data

	A	B	C
A	1	4/7	3/7
B	8/11	1	4/5
C	8/11	1	9/11

didates. Now, the number of first and second ranks obtained by the four candidates is shown in Table 3.

Table 3
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	7	1
B	4	7
C	0	6
D	3	0

When Model (1) is solved for each of the four candidates, the new scores are the following:

$$Z^*_{\mathrm{A}} = 1, \qquad Z^*_{\mathrm{B}} = 1, \qquad Z^*_{\mathrm{C}} = 6/11, \qquad Z^*_{\mathrm{D}} = 3/7.$$

Consequently, candidates A and B keep on being efficient. Table 4 shows the aggressive cross-evaluation matrix corresponding to the Table 3 data.

Table 4
Aggressive cross-evaluation matrix for the Table 3 data

	A	B	C	D
A	1	4/7	0	2/5
B	8/11	1	2/5	3/11
C	8/11	1	6/11	3/11
D	1	4/7	0	3/7

It is easy to check that the candidates' global scores are now the following:

$$\Theta_{\mathrm{A}} = 0.882 \qquad \Theta_{\mathrm{B}} = 0.757, \qquad \Theta_{\mathrm{C}} = 0.186, \qquad \Theta_{\mathrm{D}} = 0.349.$$

Therefore, candidate A is the new winner.

On the other hand, the cross-evaluation method has another shortcoming that is more serious than the previous one: a winning candidate can lose when he/she wins some j-th place ranks from inefficient candidates or, equivalently, a losing candidate can win when he/she loses some j-th place ranks to inefficient candidates. Consequently, this method is not monotonic. This fact is illustrated in the following example.

Example 2. Consider again Table 1, where candidates A and B are efficient. We have seen that the winner is candidate B. Suppose now that this candidate wins six second ranks from candidate C. In this case, Table 5 shows the number of first and second ranks obtained by each candidate.

Table 5
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	7	1
B	4	13
C	3	0

When Model (1) is solved for each candidate, we obtain the following scores:

$$Z_{\mathrm{A}}^{*} = 1, \qquad Z_{\mathrm{B}}^{*} = 1, \qquad Z_{\mathrm{C}}^{*} = 3/7,$$

i.e., candidates A and B keep on being efficient. The aggressive cross-evaluation matrix corresponding to the Table 5 data is shown in Table 6.

Table 6
Aggressive cross-evaluation matrix for the Table 5 data

	A	B	C
A	1	4/7	12/29
B	8/17	1	3/17
C	1	4/7	3/7

The winner is now candidate A, because the candidates' global scores are the following:

$$\Theta_{\mathrm{A}} = 0.787, \qquad \Theta_{\mathrm{B}} = 0.743, \qquad \Theta_{\mathrm{C}} = 0.321.$$

3 Hashimoto's model

Hashimoto [6] proposes to apply the DEA exclusion method (see Andersen and Petersen [1]) to Cook and Kress's model. This methodology enables an efficient candidate to achieve a score greater than one by removing the constraint relative to the aforementioned candidate in the formulation of Model (1). In this way, ties for first place can be broken. Moreover, Hashimoto [6] considers $d(j,\varepsilon) = \varepsilon$ for all $j \in \{1,\ldots,k\}$, with ε small enough to guarantee a decreasing sequence of weights and to avoid the solution of the model depending on the discrimination intensity functions. On the other hand, he adds new constraints to the model to assure a convex sequence of weights (for more on this, see Stein *et al.* [10]). The model proposed by this author is:

$$\begin{aligned} \widetilde{Z}_o^* = \max\ & \sum_{j=1}^{k} w_j v_{oj}, \\ \text{s.t.}\ & \sum_{j=1}^{k} w_j v_{ij} \leq 1, \quad i = 1,\ldots,m,\ i \neq o \\ & w_j - w_{j+1} \geq \varepsilon, \quad j = 1,\ldots,k-1, \\ & w_k \geq \varepsilon, \\ & w_j - 2w_{j+1} + w_{j+2} \geq 0, \quad j = 1,\ldots,k-2, \end{aligned} \tag{4}$$

where ε is a positive non-archimedian infinitesimal.

Hashimoto's model is useful to discriminate efficient candidates but it has the same drawback as the method analyzed in the previous section: the number of first, second, ..., k-th ranks obtained by inefficient candidates may change the order of efficient candidates (this fact has been pointed out, though not justified, by Obata and Ishii [8]). The next example illustrates this situation.

Example 3. Table 7 shows the number of first and second ranks obtained by five candidates.

When Model (4) is solved for each of the five candidates, the following scores are obtained:

$$\widetilde{Z}_{\mathrm{A}}^* = 1.560, \quad \widetilde{Z}_{\mathrm{B}}^* = 1.625, \quad \widetilde{Z}_{\mathrm{C}}^* = 0.5, \quad \widetilde{Z}_{\mathrm{D}}^* = 0.615, \quad \widetilde{Z}_{\mathrm{E}}^* = 0.077.$$

So, candidate B is the winner. Let us assume that candidate E loses one second rank to candidate D without any variation to the other candidates

Table 7
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	8	0
B	5	8
C	4	0
D	0	8
E	0	1

(see Table 8).

Table 8
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	8	0
B	5	8
C	4	0
D	0	9

The winner is now candidate A, since the scores obtained by the candidates are the following:

$$\widetilde{Z}^*_{\mathrm{A}} = 1.560, \quad \widetilde{Z}^*_{\mathrm{B}} = 1.514, \quad \widetilde{Z}^*_{\mathrm{C}} = 0.5, \quad \widetilde{Z}^*_{\mathrm{D}} = 0.692.$$

4 Noguchi, Ogawa and Ishii's model

Noguchi *et al.* [7] criticize the choice of discrimination intensity functions in Green *et al.*'s model. In this model, the weight assigned to a certain rank may be zero and, consequently, the votes granted to that rank are not considered. Furthermore, the weights corresponding to two different ranks may be equal and, therefore, the rank votes lose their meaning.

To avoid the previous drawbacks, Noguchi *et al.* [7] propose changing, in Green *et al.*'s model, the conditions relatives to the weights for one of the two following set of constraints:

(1) $w_j - w_{j+1} \geq \varepsilon$ for all $j \in \{1, \ldots, k-1\}$ and $w_k \geq \varepsilon$, with ε small enough (Noguchi *et al.*'s weak ordering).

(2) $w_1 \geq 2w_2 \geq 3w_3 \geq \cdots \geq kw_k$ and $w_k \geq \varepsilon = \frac{2}{nk(k+1)}$, where n is the number of voters (Noguchi *et al.*'s strong ordering).

Besides the previous conditions on the scoring vectors, Noguchi *et al.* [7] introduce two other modifications in the model of Green *et al.* [5]. On the one hand, in the cross-evaluation matrix each candidate utilizes the same scoring vector to evaluate each of the remaining candidates (without specifying that weights must be chosen when Model (1) has multiple solutions). On the other hand, the geometric mean of the scores obtained by each candidate is used to calculate the global scores.

However, Noguchi *et al.*'s models maintain the problems of Green *et al.*'s model. Thus, in relation to Noguchi *et al.*'s model with weak ordering, Examples 1 and 2 can be used, respectively, to show that the number of first, second, ..., k-th ranks obtained by inefficient candidates may change the order of efficient candidates and that this procedure is not monotonic.

The following examples show these drawbacks when Noguchi *et al.*'s model with strong ordering is used.

Example 4. Table 9 shows the number of first and second ranks obtained by four candidates.

Table 9
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	6	1
B	4	7
C	3	6
D	3	2

If we substitute the constraints about the weights by $w_1 \geq 2w_2$ and $w_2 \geq 1/48$ in Model (1), and solve it for each of the four candidates, the following scores are obtained:

$$Z^*_{\mathrm{A}} = 1, \qquad Z^*_{\mathrm{B}} = 1, \qquad Z^*_{\mathrm{C}} = 4/5, \qquad Z^*_{\mathrm{D}} = 11/19.$$

Therefore, candidates A and B are efficient. Table 10 shows the cross-evaluation matrix corresponding to the Table 9 data and the weights used by each candidate to evaluate the remaining ones.

If we use the geometric mean to obtain the scores for each candidate, the

Table 10
Cross-evaluation matrix for the Table 9 data

	w_1	w_2	A	B	C	D
A	47/288	1/48	1	115/144	59/96	17/32
B	2/15	1/15	13/15	1	4/5	8/15
C	2/15	1/15	13/15	1	4/5	8/15
D	3/19	1/19	1	1	15/19	11/19

global scores are:

$$\widetilde{\Theta}_{\mathrm{A}} = 0.931, \qquad \widetilde{\Theta}_{\mathrm{B}} = 0.945, \qquad \widetilde{\Theta}_{\mathrm{C}} = 0.746, \qquad \widetilde{\Theta}_{\mathrm{D}} = 0.544.$$

So, candidate B is the winner. Let us assume that candidates D loses a first rank to candidate C. Now, the number of first and second ranks obtained by each candidates is shown in Table 11.

Table 11
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	6	1
B	4	7
C	4	6
D	2	2

When Model (1), with the changes mentioned above, is solved for each of the four candidates, the new scores are the following:

$$Z^*_{\mathrm{A}} = 1, \qquad Z^*_{\mathrm{B}} = 1, \qquad Z^*_{\mathrm{C}} = 18/19, \qquad Z^*_{\mathrm{D}} = 8/19.$$

Consequently, candidates A and B keep on being efficient. In Table 12 we show the cross-evaluation matrix corresponding to the Table 11 data and the weights used in the evaluation.

It is easy to check that the order and global scores (using geometric mean) of the candidates are now the following:

$$\widetilde{\Theta}_{\mathrm{A}} = 0.965, \qquad \widetilde{\Theta}_{\mathrm{B}} = 0.945, \qquad \widetilde{\Theta}_{\mathrm{C}} = 0.898, \qquad \widetilde{\Theta}_{\mathrm{D}} = 0.402.$$

Therefore, candidate A is the new winner.

Table 12
Cross-evaluation matrix for the Table 11 data

	w_1	w_2	A	B	C	D
A	47/288	1/48	1	115/144	7/9	53/144
B	2/15	1/15	13/15	1	14/15	2/5
C	3/19	1/19	1	1	18/19	8/19
D	3/19	1/19	1	1	18/19	8/19

Example 5. Consider again Table 9. We have seen that the winner is candidate B. Suppose now that this candidate wins two second ranks from candidate D. In this case, Table 13 shows the number of first and second ranks obtained by each candidate.

Table 13
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	6	1
B	4	9
C	3	6
D	3	0

When Model (1), with the changes mentioned above, is solved for each of the four candidates, the new scores are the following:

$$Z^*_{\mathrm{A}} = 1, \qquad Z^*_{\mathrm{B}} = 1, \qquad Z^*_{\mathrm{C}} = 18/25, \qquad Z^*_{\mathrm{D}} = 47/96.$$

Candidates A and B remain efficient. The cross-evaluation matrix corresponding to the Table 13 data and the weights used in the evaluation are shown in Table 14.

Table 14
Cross-evaluation matrix for the Table 13 data

	w_1	w_2	A	B	C	D
A	47/288	1/48	1	121/144	59/96	47/96
B	2/17	1/17	13/17	1	12/17	6/17
C	4/25	1/25	1	1	18/25	12/25
D	47/288	1/48	1	121/144	59/96	47/96

The global scores (using geometric mean) of the candidates are:

$$\widetilde{\Theta}_{\mathrm{A}} = 0.935, \qquad \widetilde{\Theta}_{\mathrm{B}} = 0.917, \qquad \widetilde{\Theta}_{\mathrm{C}} = 0.662, \qquad \widetilde{\Theta}_{\mathrm{D}} = 0.449.$$

So, the winner is A.

Finally, it is worth noting that the constraints $w_1 \geq 2w_2 \geq 3w_3 \geq \cdots \geq kw_k$ are not sufficiently justified, as we show in the sequel. For obtaining these relationships among the weights, the authors impose the following condition:

$$w_1 - w_2 > \cdots > w_{j-1} - w_j > w_j - w_{j+1} > \cdots > w_{k-1} - w_k > 0,$$

which is very controversial because it prevents Borda's weights being used. Moreover, to guarantee this condition, the authors suppose that

$$w_{j-1} - w_j \geq w_j - \frac{j-2}{j-1} w_{j+1}.$$

However, the inequality $w_{j-1} - w_j > w_j - w_{j+1}$ is also verified if, for example, we impose the more general condition

$$w_{j-1} - w_j \geq w_j - \alpha_{j-1} w_{j+1},$$

where $\alpha_{j-1} \in [0, 1)$. From this condition and taking into account that $w_j > w_{j+1}$, we obtain the following relationship between the weights:

$$w_{j-1} \geq (2 - \alpha_{j-1}) w_j.$$

Noguchi *et al.* [7] consider $\alpha_{j-1} = \frac{j-2}{j-1}$ for obtaining $(j-1)w_{j-1} \geq jw_j$, but if for instance $\alpha_{j-1} = 0$ is utilized, the condition would be $w_{j-1} \geq 2w_j$, i.e.,

$$w_1 \geq 2w_2 \geq 4w_3 \geq \cdots \geq 2^{k-1} w_k.$$

Therefore, the choice of $w_1 \geq 2w_2 \geq 3w_3 \geq \cdots \geq kw_k$ is somewhat arbitrary.

5 Obata and Ishii's model

Obata and Ishii [8] consider that, in order to compare the maximum score obtained by each candidate, it is fair to use scoring vectors of the same size. So, they suggest normalizing the most favorable scoring vectors for each candidate.

The model proposed by these authors is the following:

$$
\begin{aligned}
\widehat{Z}_o^* \; = \; \max \; & \sum_{j=1}^{k} \widehat{w}_j v_{oj}, \\
\text{s.t.} \quad & \sum_{j=1}^{k} w_j v_{oj} = 1, \\
& \sum_{j=1}^{k} w_j v_{ij} \leq 1, \quad i = 1, \ldots, m, \; i \neq o, \\
& w_j - w_{j+1} \geq d(j, \varepsilon), \quad j = 1, \ldots, k-1, \\
& w_k \geq d(k, \varepsilon), \\
& \widehat{w} = \alpha w, \quad \alpha > 0, \\
& \|\widehat{w}\| = 1,
\end{aligned}
\tag{5}
$$

where $\|\cdot\|$ is a certain norm. Obata and Ishii [8] demonstrate that the previous model is equivalent to the following one:

$$
\begin{aligned}
\widehat{Z}_o^* \; = \; \max \; & \frac{1}{\|w\|}, \\
\text{s.t.} \quad & \sum_{j=1}^{k} w_j v_{oj} = 1, \\
& \sum_{j=1}^{k} w_j v_{ij} \leq 1, \quad i = 1, \ldots, m, \; i \neq o, \qquad \text{(a)} \\
& w_j - w_{j+1} \geq d(j, \varepsilon), \quad j = 1, \ldots, k-1, \\
& w_k \geq d(k, \varepsilon).
\end{aligned}
\tag{6}
$$

Obviously, if candidate A_o is inefficient, problem (6) has no feasible solution. Furthermore, it is easy to prove that the solution of problem (6) never changes if constraints (6a) relative to inefficient candidates are removed. Therefore, Obata and Ishii's model does not use any information about inefficient candidates and, because of this, the number of first, second, ..., k-th ranks obtained by inefficient candidates cannot change the order of efficient candidates. This is the main advantage of this model over the previous ones. However, this model presents other problems that we detail in what follows.

The first one is the choice of the norm, because the winner depends on the norm used. To illustrate this situation, we consider the following example.

Example 6. Table 15, taken from Obata and Ishii [8, p. 235], shows the number of first and second ranks obtained by seven candidates.

Table 15
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	32	10
B	28	20
C	13	36
D	20	27
E	27	19
F	30	8
G	0	30

Obata and Ishii [8] consider $d(1,\varepsilon) = d(2,\varepsilon) = 0$. Under this assumption, candidates A, B and C are efficient. When Model (6) is solved using the L_1-norm, the scores obtained by each efficient candidate are the following:

$$\widehat{Z}^*_{\mathrm{A}} = 32, \quad \widehat{Z}^*_{\mathrm{B}} = 25.714, \quad \widehat{Z}^*_{\mathrm{C}} = 24.5.$$

So, candidate A is the winner. On the other hand, if we use the L_∞-norm, the following scores are obtained:

$$\widehat{Z}^*_{\mathrm{A}} = 36, \quad \widehat{Z}^*_{\mathrm{B}} = 46.75, \quad \widehat{Z}^*_{\mathrm{C}} = 49.$$

Therefore, in this situation, the winner is candidate C.

The advantage of using L_1-norm or L_∞-norm is that Model (6) is equivalent to a linear one. Moreover, if we use these norms and the discrimination intensity functions are zero, the winner could be obtained without solving Model (6), as we show in what follows. For this, we replace standing v_{ij} by a cumulative standing $V_{ij} = \sum_{l=1}^{j} v_{il}$ (on this, see Green *et al.* [5, p. 465]). When we consider $d(j,\varepsilon)=0$ for all $j \in \{1,\ldots,k\}$, Model (6) becomes:

$$\begin{aligned} \widehat{Z}_o^* \;=\; \max \;& \frac{1}{\left\| \left(\sum\limits_{j=1}^{k} W_j, \sum\limits_{j=2}^{k} W_j, \ldots, W_k \right) \right\|}, \\ \text{s.t.} \;& \sum_{j=1}^{k} W_j V_{oj} = 1, \\ & \sum_{j=1}^{k} W_j V_{ij} \leq 1, \quad i=1,\ldots,m, \; i \neq o, \\ & W_j \geq 0, \quad j=1,\ldots,k. \end{aligned} \tag{7}$$

In the following subsections we analyze the Obata and Ishii's model using L_∞-norm and L_1-norm.

5.1 Obata and Ishii's model using L_∞-norm

When we consider L_∞-norm, Model (7) becomes:

$$\begin{aligned} \widehat{Z}_o^* \;=\; \max \;& \frac{1}{\sum\limits_{j=1}^{k} W_j} \\ \text{s.t.} \;& \sum_{j=1}^{k} W_j V_{oj} = 1, \\ & \sum_{j=1}^{k} W_j V_{ij} \leq 1, \quad i=1,\ldots,m, \; i \neq o, \\ & W_j \geq 0, \quad j=1,\ldots,k. \end{aligned} \tag{8}$$

In order to obtain the characterization of the winning candidates, we previously give two lemmas.

Lemma 7. *Consider Model (8). If A_o is an efficient candidate, then $\widehat{Z}_o^* \leq V_{ok}$.*

PROOF. This is proven by contradiction. If $\widehat{Z}_o^* > V_{ok}$, then there exists a feasible solution $(W_1^*, \ldots, W_k^*)$ such that

$$\frac{1}{\sum\limits_{j=1}^{k} W_j^*} > V_{ok},$$

i.e., $1 > \sum\limits_{j=1}^{k} W_j^* V_{ok}$. Since $V_{ok} \geq V_{oj}$ for all $j \in \{1, \ldots, k\}$, we also have $1 > \sum\limits_{j=1}^{k} W_j^* V_{oj}$, which contradicts that $(W_1^*, \ldots, W_k^*)$ is a feasible solution. □

Lemma 8. *Consider Model (8). If A_o is an efficient candidate such that $V_{ok} = \max\limits_{i=1,\ldots,m} V_{ik}$, then $\widehat{Z}_o^* = V_{ok}$.*

PROOF. By Lemma 7 we have $\widehat{Z}_o^* \leq V_{ok}$. In order to prove the equality, it is sufficient to take into account that

$$W_j = \begin{cases} \dfrac{1}{V_{ok}}, & \text{if } \ j = k, \\ \\ 0, & \text{otherwise}, \end{cases}$$

is a feasible solution of problem (8). □

Theorem 9. *Consider Model (8). An efficient candidate A_o is a winner if and only if $V_{ok} = \max\limits_{i=1,\ldots,m} V_{ik}$. Moreover, in this case, $\widehat{Z}_o^* = V_{ok}$.*

PROOF. Suppose first that A_o is an efficient candidate such that $V_{ok} = \max\limits_{i=1,\ldots,m} V_{ik}$. By Lemmas 8 and 7 we have $\widehat{Z}_o^* = V_{ok} = \max\limits_{i=1,\ldots,m} V_{ik} \geq \max\limits_{i=1,\ldots,m} \widehat{Z}_i^*$. Therefore, A_o is a winner.

For the converse, we are going to prove that if $V_{ok} < \max\limits_{i=1,\ldots,m} V_{ik} = V_{lk}$ then A_o is not a winner. By Lemmas 7 and 8 we have $\widehat{Z}_o^* \leq V_{ok} < V_{lk} = \widehat{Z}_l^*$. Consequently, A_o is not a winner. □

In the following theorem we show that the maximum value of the cumulative standings V_{ik} cannot be reached by inefficient candidates.

Theorem 10. *If A_o is a candidate such that $V_{ok} = \max\limits_{i=1,\ldots,m} V_{ik}$, then A_o is an* *efficient candidate.*

PROOF. When the cumulative standings V_{ij} are used and the discrimination intensity functions are zero, Model (1) becomes:

$$\begin{aligned} \widehat{Z}_o^* &= \max \sum_{j=1}^{k} W_j V_{oj}, \\ &\text{s.t.} \quad \sum_{j=1}^{k} W_j V_{ij} \leq 1, \quad i = 1, \ldots, m, \\ &\qquad W_j \geq 0, \quad j = 1, \ldots, k. \end{aligned} \tag{9}$$

In order to prove that A_o is an efficient candidate, it is sufficient to consider that

$$W_j = \begin{cases} \dfrac{1}{V_{ok}}, \text{ if } j = k, \\ 0, \quad \text{otherwise}, \end{cases}$$

is a feasible solution of problem (9). □

Taking into account the previous theorems, the winning candidates are those who obtain the largest sum of first, second, ..., k-th ranks. This fact is illustrated in the following example.

Example 11. Table 16 shows the cumulative standings corresponding to the Table 15 data. From Theorem 9 we obtain that C is the winner and $\widehat{Z}_C^* = 49$.

It is worth noting that the winning candidates are those who win when we use a scoring rule with $w_j = 1$ for all $j \in \{1, \ldots, k\}$ [1]. Given that the aim of the models based on DEA methodology is to avoid utilizing the same scoring vector for all candidates, L_∞ norm does not seem the best-suited to be used.

[1] Nevertheless, the total ranking including the efficient candidates behind the winner does not always coincide in both cases. For instance, if in Example 6 the number of first and second ranks obtained by candidate D had been 4 and 37 respectively, candidates A, B, C and D had been efficient. When Obata and Ishii's model is applied, the order of the candidates would have been C $\succ$ B $\succ$ D $\succ$ A. However, if the scoring rule is used, the order would have been C $\succ$ B $\succ$ A $\succ$ D.

Table 16
Cumulative standings for the Table 15 data

Candidate	V_{i1}	V_{i2}
A	32	42
B	28	48
C	13	49
D	20	47
E	27	46
F	30	38
G	0	30

5.2 Obata and Ishii's model using L_1-norm

When we consider L_1-norm, Model (7) becomes:

$$\begin{aligned} \widehat{Z}_o^* &= \max \; \frac{1}{\sum_{j=1}^{k} jW_j} \\ &\quad \text{s.t.} \;\; \sum_{j=1}^{k} W_j V_{oj} = 1, \\ &\qquad\quad \sum_{j=1}^{k} W_j V_{ij} \leq 1, \quad i = 1, \ldots, m, \; i \neq o, \\ &\qquad\quad W_j \geq 0, \quad j = 1, \ldots, k. \end{aligned} \tag{10}$$

In order to obtain the characterization of the winning candidates, we previously give three lemmas.

Lemma 12. *Consider Model (10). If A_o is an efficient candidate, then*

$$\widehat{Z}_o^* \leq \max_{j=1,\ldots,k} \frac{V_{oj}}{j}.$$

PROOF. This is proven by contradiction. Let $p \in \{1, \ldots, k\}$ such that

$$\frac{V_{op}}{p} = \max_{j=1,\ldots,k} \frac{V_{oj}}{j},$$

and suppose that there exists a feasible solution $(W_1^*, \dots, W_k^*)$ such that

$$\frac{1}{\sum_{j=1}^{k} jW_j^*} > \frac{V_{op}}{p},$$

i.e., $1 > \sum_{j=1}^{k} jW_j^* \frac{V_{op}}{p}$. Since $\frac{V_{op}}{p} \geq \frac{V_{oj}}{j}$ for all $j \in \{1, \dots, k\}$, we also have $1 > \sum_{j=1}^{k} W_j^* V_{oj}$, which contradicts that $(W_1^*, \dots, W_k^*)$ is a feasible solution. □

Lemma 13. *Consider Model (10). If A_o is an efficient candidate such that $V_{op} = \max_{i=1,\dots,m} V_{ip}$ for some $p \in \{1, \dots, k\}$, then*

$$\widehat{Z}_o^* \geq \frac{V_{op}}{p}.$$

PROOF. It is sufficient to consider that

$$W_j = \begin{cases} \frac{1}{V_{op}}, & \text{if } \ j = p, \\ 0, & \text{otherwise}, \end{cases}$$

is a feasible solution of problem (10). □

Lemma 14. *Consider Model (10). If A_o is an efficient candidate such that $V_{op} = \max_{i=1,\dots,m} V_{ip}$ and $\frac{V_{op}}{p} = \max_{j=1,\dots,k} \frac{V_{oj}}{j}$ for some $p \in \{1, \dots, k\}$, then*

$$\widehat{Z}_o^* = \frac{V_{op}}{p}.$$

PROOF. It is obvious from Lemmas 12 and 13. □

Theorem 15. *Consider Model (10). An efficient candidate A_o is a winner if and only if there exists $p \in \{1, \dots, k\}$ such that*

$$\frac{V_{op}}{p} = \max_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{V_{ij}}{j}.$$

Moreover, in this case, $\widehat{Z}_o^* = \dfrac{V_{op}}{p}$.

PROOF. Suppose first that A_o is an efficient candidate such that there exists $p \in \{1, \dots, k\}$ with

$$\frac{V_{op}}{p} = \max_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{V_{ij}}{j}.$$

By Lemmas 14 and 12 we have

$$\widehat{Z}_o^* = \frac{V_{op}}{p} = \max_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{V_{ij}}{j} \geq \max_{i=1,\dots,m} \widehat{Z}_i^*.$$

Therefore, A_o is a winner.

For the converse, we are going to prove that if for all $p \in \{1, \dots, k\}$

$$\frac{V_{op}}{p} < \max_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{V_{ij}}{j} = \frac{V_{lq}}{q},$$

then A_o is not a winner. By Lemmas 12 and 14 we have

$$\widehat{Z}_o^* \leq \max_{j=1,\dots,k} \frac{V_{oj}}{j} < \frac{V_{lq}}{q} = \widehat{Z}_l^*.$$

Consequently, A_o is not a winner. □

In a similar way to that of Theorem 10, it is easy to check that if there exists $p \in \{1, \dots, k\}$ such that $V_{op} = \max\limits_{i=1,\dots,m} V_{ip}$, then A_o is an efficient candidate. Therefore, the candidates such that

$$\frac{V_{op}}{p} = \max_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{V_{ij}}{j}$$

for some $p \in \{1, \dots, k\}$ are always efficient.

The characterization result given in Theorem 15 is illustrated in the following example.

Example 16. Table 17 shows the average of the standings corresponding to the Table 15 data. From Theorem 15 we obtain that A is the winner and $\widehat{Z}_A^* = 32$.

Table 17
Average of the standings for the Table 15 data

Candidate	$V_{i1}/1$	$V_{i2}/2$
A	32	21
B	28	24
C	13	24.5
D	20	23.5
E	27	23
F	30	19
G	0	15

It is worth emphasizing that Theorem 15 provides a necessary condition for a candidate to be a winner: the sum of votes obtained by him/her up to a certain rank must be equal to or larger than the corresponding sum for the remaining candidates. This condition can be very restrictive, as we show in the following example.

Example 17. Table 18 shows the number of first and second ranks obtained by three candidates, where v_1 and v_2 satisfy the constraints $v_1 \leq 101$ and $202 - 2v_1 \leq v_2 \leq 202 - v_1$, so that the three candidates are efficient.

Table 18
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	101	0
B	v_1	v_2
C	0	202

Moreover, we suppose that there exist inefficient candidates not considered in Table 18, so that the sum of the first ranks coincides with the sum of the second ones. If $v_1 \leq 100$ and $v_2 \leq 201 - v_1$, candidates A and C are the winners when Obata and Ishii's method is applied. Therefore, with the constraints given, it is impossible to find v_1 and v_2 values so that candidate B may be the winner. This fact originates that the result may be considered unfair from candidate B's point of view. To illustrate this situation, we consider $v_1 = 100$ and $v_2 = 101$ (see Table 19). On the one hand, since candidate A defeats B, it seems that one first rank is more valued than 101 second ones; nevertheless, as candidate C also defeats B, it seems that 101 second ranks are more valued that 100 first ones.

Table 19
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	101	0
B	100	101
C	0	202

Likewise, although the purpose of this method is to discriminate efficient candidates, when the discrimination intensity functions are zero, this procedure does not manage to discriminate in some situations in which the choice of the winning candidate should not be questioned. The following example illustrates this situation.

Example 18. Let us consider Table 20, which shows the number of first and second ranks obtained by five candidates.

Table 20
First and second ranks obtained by each candidate

Candidate	First rank	Second rank
A	100	100
B	100	0
C	0	200
D	50	0
E	50	0

Candidates A, B and C are efficient. When Model (10) is applied, the scores obtained by each efficient candidate are:

$$\widehat{Z}^*_{\mathrm{A}} = \widehat{Z}^*_{\mathrm{B}} = \widehat{Z}^*_{\mathrm{C}} = 100.$$

The three candidates tie. However, it seems logical to think that candidate A should be the winner: he/she has obtained the same number of the first ranks and 100 second ones more than candidate B, and he/she has obtained 100 first ranks more than candidate C that should be more valued than the 100 additional second ranks that candidate C has.

It is worth noting that this shortcoming can be overcome by considering $d(j,\varepsilon) = \varepsilon$ for all $j \in \{1,\ldots,k\}$, with ε small enough. In this way, the sequence of weights is decreasing and the problem of the choice of $d(j,\varepsilon)$ is avoided.

As a final comment, we point out that Foroughi and Tamiz [4] have extended Obata and Ishii's model to rank inefficient as well as efficient candidates. However, this extension continues having the same problems.

6 Conclusions

In this paper we have analyzed the principal methods proposed in the literature to discriminate efficient candidates. The main conclusion of our study is that none of the proposed procedures is fully convincing. On the one hand, it is difficult to approve a non-monotonic method (Green *et al.*'s model and Noguchi *et al.*'s models) or a method in which the number of first, second, $\ldots$, k-th ranks obtained by inefficient candidates may change the order of efficient candidates (Green *et al.*'s model, Hashimoto's model and Noguchi *et al.*'s models). On the other hand, Obata and Ishii's model avoids the previous problems but it presents other drawbacks. In this model it is necessary to determine the norm and the discrimination intensity functions to use. If these functions are zero and the L_∞-norm is used, the winning candidate coincides with the one obtained by means of a scoring rule. If L_∞-norm is replaced by L_1-norm, the outcome could be considered unfair by some candidates.

Acknowledgements

This work is partially financed by the Junta de Castilla y León (Consejería de Educación y Cultura, Projects VA040A05 and VA099/04), ERDF and Spanish Ministerio de Educación y Ciencia (Projects SEJ2006-04267 and MTM2005-06534). The authors are grateful to two anonymous referees for valuable comments.

References

[1] P. Andersen, N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis, Management Science 39 (10) (1993) 1261–1294.

[2] S.J. Brams, P.C. Fishburn, Voting procedures, in: K.J. Arrow, A.K. Sen, K. Suzumura (Eds.), Handbook of Social Choice and Welfare, Vol. 1, Elsevier, Amsterdam, 173–236, 2002.

[3] W.D. Cook, M. Kress, A data envelopment model for aggregating preference rankings, Management Science 36 (11) (1990) 1302–1310.

[4] A.A. Foroughi, M. Tamiz, An effective total ranking model for a ranked voting system, Omega 33 (2005) 491–496.

[5] R.H. Green, J.R. Doyle, W.D. Cook, Preference voting and project ranking using DEA and cross-evaluation, European Journal of Operational Research 90 (3) (1996) 461–472.

[6] A. Hashimoto, A ranked voting system using a DEA/AR exclusion model: A note, European Journal of Operational Research 97 (3) (1997) 600–604.

[7] H. Noguchi, M. Ogawa, H. Ishii, The appropriate total ranking method using DEA for multiple categorized purposes, Journal of Computational and Applied Mathematics 146 (1) (2002) 155–166.

[8] T. Obata, H. Ishii, A method for discriminating efficient candidates with ranked voting data, European Journal of Operational Research 151 (1) (2003) 233–237.

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

[10] W.E. Stein, P.J. Mizzi, R.C. Pfaffenberger, A stochastic dominance analysis of ranked voting systems with scoring, European Journal of Operational Research 74 (1) (1994) 78–85.