

Ranking candidates through convex sequences of variable weights

Bonifacio Llamazares

Abstract Scoring rules are a well-known class of positional voting systems where fixed scores are assigned to the different ranks. Nevertheless, since the winners may change according to the scores used, the choice of the scoring vector is not obvious. For this reason several methods have been suggested so that each candidate may be evaluated with the most favorable scoring vector for him/her. In this paper we propose a new model that allows to use different scoring vector for each candidate and avoid some shortcomings of other methods suggested in the literature. Moreover we give a closed expression for the score obtained by each candidate and, in this way, it is possible to rank the candidates without solving the proposed model.

Keywords Scoring rules · Data envelopment analysis · Convex sequences of weights · Variable weights.

1 Introduction

A classical problem in the decision-making field is how to get a collective ranking or a winning candidate from individual rankings of a set of candidates or alternatives. One usual way to tackle this problem is to consider scoring rules for aggregating the individual rankings. In scoring rules, fixed scores are assigned to the different ranks obtained by the candidates and these ones are ordered according to the total number of points they receive. Notice that scoring rules are the best-known examples of positional voting systems (see Llamazares and Peña (2015a)), where voters rank order the candidates from best to worst and a set of winners is selected using the positions of the candidates in the voters' preference orders.

Bonifacio Llamazares
Departamento de Economía Aplicada, Instituto de Matemáticas (IMUVA), Universidad de Valladolid,
Avda. Valle de Esgueva 6, 47011 Valladolid, Spain.
Tel.: +34 983 186544
Fax: +34 983 423299
E-mail: boni@eco.uva.es

Due to their simplicity and good properties, scoring rules have received considerable attention in the literature (see, for instance, Llamazares and Peña (2015b) and references therein). Nowadays, scoring rules are used in sport competitions like the *Formula One World Championship*, the *IndyCar Series Championship* or the *Motorcycle World Championship*. Likewise, they are also used for awarding the *FIFA Ballon d'Or Award*, the *Baseball Writers Most Valuable Player Award* or the *Most Valuable Player of the National Basketball Association (MVP of the NBA)*.

However, one of the most important issues in the field of scoring rules is the choice of the scoring vector to use, because a candidate that is not the winner with the scoring vector imposed initially could be it if another one is used. For instance, the scoring vector used in the Formula One World Championship has changed several times. Thus, from 2003 to 2009, the scoring vector used for selecting the winner of the championship was $(10, 8, 6, 5, 4, 3, 2, 1)$. In 2008, the winner was Lewis Hamilton, followed by Felipe Massa. However, from 1991 to 2002, only the six first positions were considered and the scoring vector used was $(10, 6, 4, 3, 2, 1)$. If this scoring vector had been used in 2008, then the winner would have been Felipe Massa (see Llamazares and Peña (2013) for another example of this).

To avoid the previous problem, Cook and Kress (1990) introduced Data Envelopment Analysis (DEA) in this context in order to evaluate each candidate with the most favorable scoring vector for him/her. However, one important shortcoming of their model is that several candidates are often efficient, i.e., they achieve the maximum attainable score. For this reason, some procedures to discriminate efficient candidates have appeared in the literature (see, for instance, Green et al (1996), Hashimoto (1997), Noguchi et al (2002) and Obata and Ishii (2003)). Nevertheless, as it has been noticed by Llamazares and Peña (2009), some of the previous models have a serious drawback from the point of view of Social Choice Theory: the relative order between two candidates may be altered when the number of first, second, . . . , k th ranks obtained by other candidates changes, although there is not any variation in the number of first, second, . . . , k th ranks obtained by both candidates.

On the basis of the pioneering work of Cook and Kress (1990), several models have appeared in the literature to deal with this kind of problems (see, for instance, Hashimoto and Wu (2004), Contreras et al (2005), Foroughi et al (2005), Foroughi and Tamiz (2005), Wang and Chin (2007), Wang et al (2007a), Wang et al (2007b), Wang et al (2008), Wu et al (2009), Amin and Sadeghi (2010), Soltanifar et al (2010), Contreras (2011), Hosseinzadeh Lotfi and Fallahnejad (2011), Ebrahimnejad (2012), Foroughi and Aouni (2012), Hosseinzadeh Lotfi et al (2013), Llamazares and Peña (2013) and Hadi-Vencheh (2014)).

Among this great variety of models, we would like to point out those of Hashimoto (1997) and Llamazares and Peña (2013). Although Hashimoto's model has the above-described shortcoming (see Llamazares and Peña (2009)), it is very interesting because it uses the DEA super-efficiency model (see Andersen and Petersen (1993)) for breaking ties for first place. Moreover, this model also considers convex sequences of weights, which is a very natural condition in this context (see Stein et al (1994)). For their part, Llamazares and Peña (2013) avoid the above-described shortcoming by putting together in a single restriction the constraints of the candidates that are not being evaluated.

In this paper we take into account both methodologies. So, we propose and analyze a model with convex sequences of weights where we put together in a single restriction the constraints of the candidates that are not being evaluated. In this way, in our model, the relative order between two candidates cannot be altered by variations in the number of first, second, \dots , k th ranks obtained by the remaining candidates. Moreover, we also give a closed expression for the scores assigned to the candidates, and thus we can obtain the winning candidates without solving the proposed model.

The rest of the paper is organized as follows. In Section 2 we recall Cook and Kress' model and Hashimoto's model. In Section 3 we present our model and give a closed expression for the scores obtained by the candidates. Finally, Section 4 is devoted to conclusions. All proofs are in the Appendix.

2 Models of Cook and Kress (1990) and Hashimoto (1997)

Let $\mathcal{A} = \{A_1, \dots, A_m\}$ be a set of candidates and suppose that each voter selects k candidates and ranks them from top to k th place. Under the scoring rule associated with the scoring vector (w_1, \dots, w_k) , the candidate A_i receives $Z_i = \sum_{j=1}^k v_{ij}w_j$ points, where v_{ij} is the number of j th place ranks that candidate A_i occupies, and the candidates are ordered according to the score obtained. Two of the best known scoring rules are the plurality rule, where $w_1 = 1$ and $w_j = 0$ for all $j \in \{2, \dots, k\}$, and the Borda rule, where $k = m$ and $w_j = m - j$ for all $j \in \{1, \dots, m\}$.

One of the most important questions in this topic is the choice of the scoring vector to use, given that this choice could determine the winning candidate. To avoid this problem, Cook and Kress (1990) suggested to evaluate each candidate with the most favorable scoring vector for him/her. The model DEA/AR (Data Envelopment Analysis/Assurance Region) suggested by these authors was

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

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

One important shortcoming of the previous model is that several candidates are often efficient, i.e., they achieve the maximum attainable score ($Z_o^*(\varepsilon) = 1$). For this reason, Hashimoto (1997) proposed to apply the DEA super-efficiency model (see Andersen and Petersen (1993)) to Cook and Kress's model: By removing in Model (1) the constraint relative to the candidate that is being evaluated, efficient candidates can achieve scores greater than one and, in this way, ties for first place can be broken. Moreover, Hashimoto (1997) considered $d(j, \varepsilon) = \varepsilon$ for all $j \in \{1, \dots, k\}$, 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 added new constraints to the model to assure a convex sequence of weights (for more on this, see Stein et al (1994)); that is, $w_j - w_{j+1} \geq w_{j+1} - w_{j+2}$ for $j = 1, \dots, k-2$ (or equivalently, $w_j - 2w_{j+1} + w_{j+2} \geq 0$ for $j = 1, \dots, k-2$). So, the model proposed by this author was

$$\begin{aligned} \tilde{Z}_o^*(\varepsilon) = \max \quad & \sum_{j=1}^k v_{oj} w_j, \\ \text{s.t.} \quad & \sum_{j=1}^k v_{ij} w_j \leq 1, \quad i = 1, \dots, m, i \neq o \end{aligned} \quad (2.1)$$

$$w_j - w_{j+1} \geq \varepsilon, \quad j = 1, \dots, k-1, \quad (2.2)$$

$$w_k \geq \varepsilon, \quad (2.3)$$

$$w_j - 2w_{j+1} + w_{j+2} \geq 0, \quad j = 1, \dots, k-2, \quad (2.4)$$

where ε is a positive non-archimedean infinitesimal.

Although Hashimoto's model allows to discriminate efficient candidates, it has an important drawback: the number of first, second, \dots , k th ranks obtained by inefficient candidates may change the order of efficient candidates (see an example of this shortcoming in Llamazares and Peña (2009)). For this reason in the next section we propose a model that avoids the above drawback. This model is based on Hashimoto's model and in the methodology suggested by Llamazares and Peña (2013).

3 Our model

Consider Hashimoto's model. Notice that constraints (2.2) are equivalent to $w_{k-1} - w_k \geq \varepsilon$ due to the convex sequence of weights. Moreover, in this restriction, instead of using the parameter ε , we consider the weight w_k , which has been considered by authors such as Stein et al (1994) and Contreras et al (2005).

Now, to avoid that the positions obtained by inefficient candidates may change the order of efficient candidates, we put together in a single restriction the constraints of the candidates that are not being evaluated (see Llamazares and Peña (2013)). So, we replace the constraints (2.1) by their sum. Since

$$\sum_{\substack{i=1 \\ i \neq o}}^m \sum_{j=1}^k v_{ij} w_j = \sum_{j=1}^k w_j \sum_{\substack{i=1 \\ i \neq o}}^m v_{ij} = \sum_{j=1}^k w_j (n - v_{oj}),$$

where n is the number of voters, we get the constraint

$$\sum_{j=1}^k (n - v_{oj}) w_j \leq m - 1.$$

Therefore, our model can be expressed as follows:

$$\begin{aligned}
\widehat{Z}_o^*(\varepsilon) = \max \quad & \sum_{j=1}^k v_{oj} w_j, \\
\text{s.t.} \quad & \sum_{j=1}^k (n - v_{oj}) w_j \leq m - 1, \\
& w_{k-1} - w_k \geq w_k, \\
& w_k \geq \varepsilon, \\
& w_j - 2w_{j+1} + w_{j+2} \geq 0, \quad j = 1, \dots, k-2,
\end{aligned} \tag{3}$$

where we maximize the score of each candidate under the assumption that the total score of the remaining candidates is less than or equal to the number of candidates minus 1. Moreover, we also suppose convex sequences of weights.

In order to make easier the analysis of this model, in the following lemma we give an alternative representation of it. In this way, we get an equivalent model where we have replaced the convexity restriction on the variables by nonnegativity.

Lemma 1 *Model (3) can be expressed as*

$$\begin{aligned}
\widehat{Z}_o^*(\varepsilon) = \max \quad & \sum_{j=1}^k V_{oj} W_j + \varepsilon V_{ok}, \\
\text{s.t.} \quad & \sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j \leq \delta_o(\varepsilon), \\
& W_j \geq 0, \quad j = 1, \dots, k,
\end{aligned} \tag{4}$$

where

$$\begin{aligned}
W_j &= w_j - 2w_{j+1} + w_{j+2}, \quad \text{for all } j \in \{1, \dots, k-2\}, \\
W_{k-1} &= w_{k-1} - 2w_k, \\
W_k &= w_k - \varepsilon, \\
V_{oj} &= \sum_{l=1}^j (j+1-l)v_{ol}, \quad \text{for all } j \in \{1, \dots, k\}, \\
\delta_o(\varepsilon) &= (m-1) - \left(\frac{nk(k+1)}{2} - V_{ok} \right) \varepsilon.
\end{aligned}$$

Notice that

$$V_{ok} = kv_{o1} + (k-1)v_{o2} + \dots + v_{ok} \leq kn \leq \frac{nk(k+1)}{2}.$$

Therefore, the function $\delta_o(\varepsilon)$ is nonincreasing in ε and, consequently, the feasible set of Model (4) does not increase when the value of ε increases. Hence $\widehat{Z}_o^*(\varepsilon)$ is a nonincreasing function; that is, if $\varepsilon_1 > \varepsilon_2$, then $\widehat{Z}_o^*(\varepsilon_1) \leq \widehat{Z}_o^*(\varepsilon_2)$.

In order to ensure that Model (4) is feasible, we need to impose the condition $\delta_o(\varepsilon) \geq 0$. Given that the feasible set of the previous model depends on the evaluated

candidate, the condition $\min_{o=1,\dots,m} \delta_o(\varepsilon) \geq 0$ is necessary to guarantee that all the feasible sets are non-empty.

On the other hand, it is worth noting that if a candidate A_o gets all the first ranks, then he/she is the winner. Given that $V_{oj} = jv_{o1} + (j-1)v_{o2} + \dots + v_{oj}$, when $v_{o1} = n$ (and, consequently, $v_{o2} = \dots = v_{ok} = 0$), we have $V_{oj} = jn$ for all $j \in \{1, \dots, k\}$. Therefore, for this candidate, the feasible set of Model (4) is

$$S = \left\{ (W_1, \dots, W_k) \in \mathbb{R}_+^k \mid \sum_{j=2}^k \frac{nj(j-1)}{2} W_j \leq \delta_o(\varepsilon) \right\},$$

and Model (4) is unbounded. Consequently, candidate A_o is the winner.

In the following theorem we give the optimal value of this program for the remaining cases.

Theorem 1 Consider Model (4) and let $V_{o1} < n$. Then $\widehat{Z}_o^*(\varepsilon) = \delta_o(\varepsilon)V_o^* + \varepsilon V_{ok}$, where

$$V_o^* = \max_{j=1,\dots,k} V_o^j \quad \text{and} \quad V_o^j = \frac{V_{oj}}{\frac{nj(j+1)}{2} - V_{oj}}.$$

In the following subsections we analyze the behavior of our model according to whether the parameter ε is null or not.

3.1 The case $\varepsilon = 0$

When $\varepsilon = 0$, Model (4) can be written as

$$\begin{aligned} \widehat{Z}_o^* &= \max \sum_{j=1}^k V_{oj} W_j, \\ \text{s.t.} \quad &\sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j \leq m-1, \\ &W_j \geq 0, \quad j = 1, \dots, k. \end{aligned} \tag{5}$$

By Theorem 1, the score obtained by the candidate A_o when $V_{o1} < n$ is

$$\widehat{Z}_o^* = (m-1)V_o^* = (m-1) \max_{j=1,\dots,k} \frac{V_{oj}}{\frac{nj(j+1)}{2} - V_{oj}}.$$

For instance, consider the example given in Table 1, taken from Cook and Kress (1990, p. 1309). There is a tie for the first place between candidates C and D, and the order of the remaining candidates is B, A, F and E. A way to break the tie between C and D will be explained in the next subsection.

It is worth noting that the order obtained through the scores \widehat{Z}_o^* is the same as obtained by using a model proposed by Contreras et al (2005, Prop. 3.4).

Table 1 Ranks obtained by each candidate and values of \widehat{Z}_o^* .

| Candidate | v_{i1} | v_{i2} | v_{i3} | v_{i4} | \widehat{Z}_o^* |
|-----------|----------|----------|----------|----------|-------------------|
| A | 3 | 3 | 4 | 3 | 0.9524 |
| B | 4 | 5 | 5 | 2 | 1.4516 |
| C | 6 | 2 | 3 | 2 | 2.1429 |
| D | 6 | 2 | 2 | 6 | 2.1429 |
| E | 0 | 4 | 3 | 4 | 0.6180 |
| F | 1 | 4 | 3 | 3 | 0.7143 |

Proposition 1 *The rank given by Model (5) is the same as the obtained by using the expression*

$$Z_o = \max_{j=1,\dots,k} \frac{V_{oj}}{j(j+1)}.$$

As Contreras et al (2005) have pointed out, the score Z_o can be interpreted as the result of evaluating the candidates by using the normalized truncated Borda rules (on this, see, Fishburn (1974)) and choosing the maximum value. Moreover, the model proposed by these authors has some interesting properties such as monotonicity, Pareto-optimality and immunity to the absolute winner paradox (see Contreras et al (2005) and Llamazares and Peña (2015a)). Therefore, Model (5) also satisfies these properties.

3.2 The case $\varepsilon > 0$

By Theorem 1, when $V_{o1} < n$ we can express the score obtained by the candidate A_o as

$$\begin{aligned} \widehat{Z}_o^*(\varepsilon) &= (m-1)V_o^* + \varepsilon \left(V_{ok} - V_o^* \left(\frac{nk(k+1)}{2} - V_{ok} \right) \right) \\ &= (m-1)V_o^* + \varepsilon \left(V_{ok}(1+V_o^*) - V_o^* \frac{nk(k+1)}{2} \right). \end{aligned} \quad (6)$$

As we can see, the graph of $\widehat{Z}_o^*(\varepsilon)$ is a straight line. Moreover, given that $\widehat{Z}_o^*(\varepsilon)$ is a nonincreasing function, the slope of this straight line is negative or null.

Notice that since

$$V_o^k = \frac{V_{ok}}{\frac{nk(k+1)}{2} - V_{ok}},$$

we get

$$V_o^k \frac{nk(k+1)}{2} = V_{ok}(1+V_o^k).$$

Therefore, when $V_o^* = V_o^k$ then

$$V_{ok}(1+V_o^*) - V_o^* \frac{nk(k+1)}{2} = 0$$

and, consequently,

$$\widehat{Z}_o^*(\varepsilon) = (m-1)V_o^k,$$

that is, the value $\widehat{Z}_o^*(\varepsilon)$ does not depend on the choice of ε .

Consider now the following example, taken from Obata and Ishii (2003) (see Table 2).

Table 2 First and second ranks obtained by each candidate.

| Candidate | v_{i1} | v_{i2} |
|-----------|----------|----------|
| A | 32 | 10 |
| B | 28 | 20 |
| C | 13 | 36 |
| D | 20 | 27 |
| E | 27 | 19 |
| F | 30 | 8 |
| G | 0 | 30 |

In this case $m = 7$, $n = 150$ and $k = 2$. Therefore

$$\widehat{Z}_o^*(\varepsilon) = 6V_o^* + \varepsilon(V_{ok}(1 + V_o^*) - 450V_o^*).$$

When we focus on candidates A and B we have

$$\widehat{Z}_A^*(\varepsilon) = \frac{96}{59} - \frac{1650}{59}\varepsilon, \quad \widehat{Z}_B^*(\varepsilon) = \frac{84}{61} - \frac{600}{61}\varepsilon.$$

Both functions appear drawn in Figure 1 (note that we have considered different scales on both axes; a hundredth on the x -axis is equal to one unit on the y -axis).

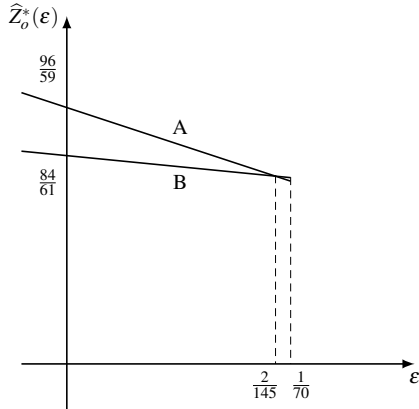


Fig. 1 Graphs of the functions $\widehat{Z}_A^*(\varepsilon)$ and $\widehat{Z}_B^*(\varepsilon)$.

As we can see in Figure 1, when we take values of ε less than $2/145$ we have $A \succ B$. However, if the values of ε are greater than $2/145$, then $B \succ A$. In Table 3

we show this behavior for two specific values of ε , $\varepsilon = 0.01$ and $\varepsilon = 1/70$. This last value is the maximum possible value for ε , that is, the maximum value for which $\min_{o=1,\dots,m} \delta_o(\varepsilon) \geq 0$.

Table 3 Values of $\delta_o(\varepsilon)$ and $\widehat{Z}_o^*(\varepsilon)$ for the candidates of Table 2.

| Candidate | $\varepsilon = 0.01$ | | $\varepsilon = 1/70$ | |
|-----------|-------------------------|--------------------------------|-------------------------|--------------------------------|
| | $\delta_o(\varepsilon)$ | $\widehat{Z}_o^*(\varepsilon)$ | $\delta_o(\varepsilon)$ | $\widehat{Z}_o^*(\varepsilon)$ |
| A | 2.24 | 1.3475 | 0.6286 | 1.2276 |
| B | 2.26 | 1.2787 | 0.6571 | 1.2365 |
| C | 2.12 | 0.9588 | 0.4571 | 0.9588 |
| D | 2.17 | 1.0496 | 0.5286 | 1.0496 |
| E | 2.23 | 1.2195 | 0.6143 | 1.1777 |
| F | 2.18 | 1.2250 | 0.5429 | 1.1071 |
| G | 1.8 | 0.7143 | 0.0000 | 0.4286 |

In the light of the previous example, it does not seem obvious how to choose a specific value of ε . One possibility would be to consider a value of ε small enough to avoid that the winner depends on the choice of ε (this is the solution proposed by some authors in their models; see, for instance, Hashimoto (1997) and Foroughi and Tamiz (2005)). If we consider in our model an infinitesimal positive value of ε , then, by expression (6), the candidates are ordered according to the score $(m-1)V_o^*$; that is, the score obtained by the candidates when $\varepsilon = 0$. If for $\varepsilon = 0$ there is a tie between two candidates, A_i and A_p , then this means that V_i^* is equal to V_p^* . Therefore,

$$1 + V_i^* = 1 + V_p^* \quad \text{and} \quad V_i^* \frac{nk(k+1)}{2} = V_p^* \frac{nk(k+1)}{2},$$

and, consequently,

$$\widehat{Z}_i^*(\varepsilon) > \widehat{Z}_p^*(\varepsilon) \Leftrightarrow V_{ik} > V_{pk}.$$

Therefore, the use of an infinitesimal positive value of ε in our model is equivalent to ranking the candidates according to the score \widehat{Z}_o^* (that is, to consider $\varepsilon = 0$) and breaking the ties between the candidates according to the value of V_{ok} . For instance, if we consider again Table 1, the tie between C and D can be broken by taken into account the values $V_{C4} = 38$ and $V_{D4} = 40$. So, in this case, $D \succ C$.

Another possibility would be to take the maximum possible value for ε (in the same spirit that in Cook and Kress' model), although in the previous example (see Figure 1), it does seem the best choice. On the other hand, this solution has a serious shortcoming: the order between two candidates may depend on the ranks obtained by other candidates. For instance, suppose that candidates C, D and G obtain the ranks shown in Table 4, which are different from those shown in Table 2. Now, the maximum possible value for ε is $1/73$, and, with this value, $A \succ B$ ($\widehat{Z}_A^*(\varepsilon)$ and $\widehat{Z}_B^*(\varepsilon)$ are still the functions of Figure 1 and $1/73 < 2/145$; see also Table 4).

To avoid taking a fixed value of ε , Llamazares and Peña (2013) propose a model where the average value of the objective functions is considered (in our case, the average of the functions $\widehat{Z}_o^*(\varepsilon)$). In the sequel we follow this methodology. Moreover,

Table 4 New ranks and values of $\delta_o(\varepsilon)$ and $\widehat{Z}_o^*(\varepsilon)$ for $\varepsilon = 1/73$.

| Candidate | v_{i1} | v_{i2} | δ_o | $\widehat{Z}_o^*(\varepsilon)$ |
|-----------|----------|----------|------------|--------------------------------|
| A | 32 | 10 | 0.8493 | 1.2440 |
| B | 28 | 20 | 0.8767 | 1.2423 |
| C | 13 | 46 | 0.8219 | 1.1176 |
| D | 20 | 35 | 0.8630 | 1.1784 |
| E | 27 | 19 | 0.8356 | 1.1834 |
| F | 30 | 8 | 0.7671 | 1.1233 |
| G | 0 | 12 | 0.0000 | 0.1644 |

for each candidate A_o we only consider the constraint $\delta_o(\varepsilon) \geq 0$ instead of the restriction $\min_{o=1,\dots,m} \delta_o(\varepsilon) \geq 0$. This prevents that the average of the function $\widehat{Z}_o^*(\varepsilon)$ may depend on the results obtained for the remaining candidates.

The maximum value for which the feasible set of Model (4) is not-empty is

$$\varepsilon_o^* = \sup \{ \varepsilon \geq 0 \mid \delta_o(\varepsilon) \geq 0 \}.$$

Since

$$\delta_o(\varepsilon) = (m-1) - \left(\frac{nk(k+1)}{2} - V_{ok} \right) \varepsilon,$$

we have

$$\varepsilon_o^* = \frac{m-1}{\frac{nk(k+1)}{2} - V_{ok}}.$$

Once known the value of ε_o^* , the score assigned to the candidate A_o is

$$\bar{Z}_o = \frac{1}{\varepsilon_o^*} \int_0^{\varepsilon_o^*} \widehat{Z}_o^*(\varepsilon) d\varepsilon,$$

that is, the average of the function $\widehat{Z}_o^*(\varepsilon)$. In the following theorem we show explicitly the value of \bar{Z}_o .

Theorem 2 Consider Model (4). Then

$$\bar{Z}_o = (m-1) \frac{V_o^* + V_o^k}{2}.$$

Notice that \bar{Z}_o is the average of \widehat{Z}_o^* and $(m-1)V_o^k$, and that, given two candidates A_i and A_p ,

$$\begin{aligned} V_i^k > V_p^k &\Leftrightarrow \frac{V_{ik}}{\frac{nk(k+1)}{2} - V_{ik}} > \frac{V_{pk}}{\frac{nk(k+1)}{2} - V_{pk}} \\ &\Leftrightarrow \frac{V_{ik}}{nk(k+1) - 2V_{ik}} > \frac{V_{pk}}{nk(k+1) - 2V_{pk}} \\ &\Leftrightarrow V_{ik} (nk(k+1) - 2V_{pk}) > V_{pk} (nk(k+1) - 2V_{ik}) \\ &\Leftrightarrow V_{ik} > V_{pk}. \end{aligned}$$

Therefore, if several candidates have the same score \widehat{Z}_o^* , then ranking these candidates according to the values \overline{Z}_o gives the same result as breaking the ties with the values V_{ok} .

It is also worth noting that, in general, \overline{Z}_o and \widehat{Z}_o^* provide different ranks. For instance, consider Table 5. The rank obtained with \widehat{Z}_o^* is

$$A \succ F \succ B \succ E \succ D \succ C \succ G.$$

while the rank obtained with \overline{Z}_o is

$$A \succ B \succ F \succ E \succ D \succ C \succ G.$$

Table 5 Values of \widehat{Z}_o^* and \overline{Z}_o for the candidates of Table 2.

| Candidate | \widehat{Z}_o^* | \overline{Z}_o |
|-----------|-------------------|------------------|
| A | 1.6271 | 1.4040 |
| B | 1.3770 | 1.2982 |
| C | 0.9588 | 0.9588 |
| D | 1.0496 | 1.0496 |
| E | 1.3171 | 1.2394 |
| F | 1.5000 | 1.2840 |
| G | 0.4286 | 0.4286 |

4 Concluding remarks

In the last years, increasing attention has been devoted to the study of ranked voting systems where each candidate is evaluated with the most favorable scoring vector for him/her. However, some of them have an important shortcoming: the relative order between two candidates may be altered when the number of first, second, \dots , k th ranks obtained by other candidates changes, although there is not any variation in the number of first, second, \dots , k th ranks obtained by both candidates. Likewise, some models are formulated using functions depending on a parameter ε , and the order between two candidates may change according to the value of ε .

In this paper we have proposed a model that avoids the above problems and that considers convex sequences of weights. Moreover, an important advantage of our model against other methods suggested in the literature is that we give a closed expression for the scores assigned to the candidates. Thus, we can obtain the order of the candidates without solving the proposed model and, in this way, it could be easy to implement in some contexts (for instance, in academic environments).

Acknowledgements

The author is very grateful to two anonymous referees for valuable suggestions, comments and references. The financial support from the Ministerio español de Economía

y Competitividad (Project ECO2012-32178) and the Junta de Castilla y León (Consejería de Educación, Project VA066U13) is gratefully acknowledged.

A Appendix

Proof of Lemma 1. Given Model (3), we consider the following change of variables:

$$\begin{cases} s_j = w_j - w_{j+1}, & \text{for all } j \in \{1, \dots, k-1\}, \\ s_k = w_k. \end{cases}$$

It is easy to check that $w_j = \sum_{p=j}^k s_p$ for all $j \in \{1, \dots, k\}$ and that the set of constraints

$$\begin{cases} w_j - w_{j+1} \geq w_{j+1} - w_{j+2}, & j = 1, \dots, k-2, \\ w_{k-1} - w_k \geq w_k, \\ w_k \geq \varepsilon. \end{cases}$$

is equivalent to the set

$$\begin{cases} s_j \geq s_{j+1}, & j = 1, \dots, k-1, \\ s_k \geq \varepsilon. \end{cases}$$

Let us make yet another change of variables:

$$\begin{cases} W_j = s_j - s_{j+1}, & \text{for all } j \in \{1, \dots, k-1\}, \\ W_k = s_k - \varepsilon. \end{cases}$$

From the previous relationships, variables W_j can be expressed as functions of variables w_j in the following way:

$$\begin{cases} W_j = s_j - s_{j+1} = w_j - 2w_{j+1} + w_{j+2}, & \text{for all } j \in \{1, \dots, k-2\}, \\ W_{k-1} = s_{k-1} - s_k = w_{k-1} - 2w_k, \\ W_k = w_k - \varepsilon. \end{cases}$$

Furthermore, it is easy to see that $s_j = \sum_{l=j}^k W_l + \varepsilon$ for all $j \in \{1, \dots, k\}$. Therefore,

$$\begin{aligned} w_j &= \sum_{p=j}^k s_p = \sum_{p=j}^k \left(\sum_{l=p}^k W_l + \varepsilon \right) = \sum_{\substack{j \leq p \leq k \\ p \leq l \leq k}} W_l + \sum_{p=j}^k \varepsilon = \sum_{j \leq p \leq l \leq k} W_l + (k+1-j)\varepsilon \\ &= \sum_{l=j}^k \sum_{p=j}^l W_l + (k+1-j)\varepsilon = \sum_{l=j}^k (l+1-j)W_l + (k+1-j)\varepsilon. \end{aligned}$$

Next we write the remaining expressions of Model (3) as functions of the variables W_j :

$$\begin{aligned}
\sum_{j=1}^k v_{oj}w_j &= \sum_{j=1}^k v_{oj} \left(\sum_{l=j}^k (l+1-j)W_l + (k+1-j)\varepsilon \right) \\
&= \sum_{\substack{1 \leq j \leq k \\ j \leq l \leq k}} (l+1-j)v_{oj}W_l + \varepsilon \sum_{j=1}^k (k+1-j)v_{oj} \\
&= \sum_{1 \leq j \leq l \leq k} (l+1-j)v_{oj}W_l + \varepsilon \sum_{j=1}^k (k+1-j)v_{oj} \\
&= \sum_{\substack{1 \leq l \leq k \\ 1 \leq j \leq l}} (l+1-j)v_{oj}W_l + \varepsilon \sum_{j=1}^k (k+1-j)v_{oj} \\
&= \sum_{l=1}^k W_l \left(\sum_{j=1}^l (l+1-j)v_{oj} \right) + \varepsilon \sum_{j=1}^k (k+1-j)v_{oj} \\
&= \sum_{j=1}^k W_j \left(\sum_{l=1}^j (j+1-l)v_{ol} \right) + \varepsilon \sum_{l=1}^k (k+1-l)v_{ol},
\end{aligned}$$

where the last equality is obtained by changing the role of j and l . If we denote $\sum_{l=1}^j (j+1-l)v_{ol}$ by V_{oj} for all $j \in \{1, \dots, k\}$, then we have

$$\sum_{j=1}^k v_{oj}w_j = \sum_{j=1}^k V_{oj}W_j + \varepsilon V_{ok}.$$

Analogously,

$$\sum_{j=1}^k (n - v_{oj})w_j = n \sum_{j=1}^k w_j - \sum_{j=1}^k v_{oj}w_j = n \sum_{j=1}^k w_j - \sum_{j=1}^k V_{oj}W_j - \varepsilon V_{ok}.$$

Given that

$$\begin{aligned}
\sum_{j=1}^k w_j &= \sum_{j=1}^k \left(\sum_{l=j}^k (l+1-j)W_l + (k+1-j)\varepsilon \right) \\
&= \sum_{\substack{1 \leq j \leq k \\ j \leq l \leq k}} (l+1-j)W_l + \varepsilon \sum_{j=1}^k (k+1-j) = \sum_{1 \leq j \leq l \leq k} (l+1-j)W_l + \varepsilon \frac{k(k+1)}{2} \\
&= \sum_{\substack{1 \leq l \leq k \\ 1 \leq j \leq l}} (l+1-j)W_l + \varepsilon \frac{k(k+1)}{2} = \sum_{l=1}^k W_l \left(\sum_{j=1}^l (l+1-j) \right) + \varepsilon \frac{k(k+1)}{2} \\
&= \sum_{l=1}^k \frac{l(l+1)}{2} W_l + \varepsilon \frac{k(k+1)}{2},
\end{aligned}$$

we have

$$\begin{aligned}\sum_{j=1}^k (n - v_{oj})w_j &= \sum_{j=1}^k \frac{nj(j+1)}{2}W_j + \varepsilon \frac{nk(k+1)}{2} - \sum_{j=1}^k V_{oj}W_j - \varepsilon V_{ok} \\ &= \sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j + \left(\frac{nk(k+1)}{2} - V_{ok} \right) \varepsilon.\end{aligned}$$

Therefore, the constraint

$$\sum_{j=1}^k (n - v_{oj})w_j \leq m - 1$$

can be written as

$$\sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j + \left(\frac{nk(k+1)}{2} - V_{ok} \right) \varepsilon \leq m - 1.$$

If we consider

$$\delta_o(\varepsilon) = (m - 1) - \left(\frac{nk(k+1)}{2} - V_{ok} \right) \varepsilon,$$

then Model (3) can be expressed as

$$\begin{aligned}\widehat{Z}_o^*(\varepsilon) &= \max \sum_{j=1}^k V_{oj}W_j + \varepsilon V_{ok}, \\ \text{s.t. } &\sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j \leq \delta_o(\varepsilon), \\ &W_j \geq 0, \quad j = 1, \dots, k.\end{aligned} \quad \square$$

Proof of Theorem 1. Model (4) is equivalent to the following one:

$$\begin{aligned}\widehat{Z}'_o(\varepsilon) &= \max \sum_{j=1}^k V_{oj}W_j, \\ \text{s.t. } &\sum_{j=1}^k \left(\frac{nj(j+1)}{2} - V_{oj} \right) W_j \leq \delta_o(\varepsilon), \\ &W_j \geq 0, \quad j = 1, \dots, k.\end{aligned}$$

Moreover $\widehat{Z}_o^*(\varepsilon) = \widehat{Z}'_o(\varepsilon) + \varepsilon V_{ok}$. It is well known that if a linear program has an optimal solution, then its dual also has an optimal solution and the optimal values for both problems are equal. Therefore, it is sufficient to solve the dual of the previous problem, that is,

$$\begin{aligned}\min &\delta_o(\varepsilon)X, \\ \text{s.t. } &\left(\frac{nj(j+1)}{2} - V_{oj} \right) X \geq V_{oj}, \quad j = 1, \dots, k, \\ &X \geq 0.\end{aligned}$$

It is easy to check that the optimal solution is

$$X^* = V_o^* = \max_{j=1, \dots, k} \frac{V_{oj}}{\frac{nj(j+1)}{2} - V_{oj}}.$$

Therefore, $\widehat{Z}'_o(\varepsilon) = \delta_o(\varepsilon)V_o^*$ and $\widehat{Z}^*_o(\varepsilon) = \delta_o(\varepsilon)V_o^* + \varepsilon V_{ok}$. \square

Proof of Proposition 1. We distinguish two cases:

1. If a candidate A_o gets all the first ranks, then he/she is the winner according with Model (5). On the other hand, since $v_{o1} = n$, we have $V_{oj} = jv_{o1} + (j-1)v_{o2} + \dots + v_{oj} = jn$. Therefore,

$$Z_o = \max_{j=1, \dots, k} \frac{V_{oj}}{\frac{j(j+1)}{2}} = \max_{j=1, \dots, k} \frac{2n}{j+1} = n.$$

Let A_i be a different candidate. Then $v_{i1} = 0$, $V_{ij} = (j-1)v_{i2} + \dots + v_{ij} \leq (j-1)n$, and

$$Z_i = \max_{j=1, \dots, k} \frac{V_{ij}}{\frac{j(j+1)}{2}} \leq \max_{j=1, \dots, k} \frac{2(j-1)n}{j(j+1)} < n.$$

Therefore $Z_o > Z_i$.

2. If no candidate obtains all the first ranks, let A_o and A_i be two candidates such that $V_{o1} < n$ and $V_{i1} < n$. Given that

$$\begin{aligned} & \frac{V_{oj}}{\frac{nj(j+1)}{2} - V_{oj}} > \frac{V_{il}}{\frac{nl(l+1)}{2} - V_{il}} \\ \Leftrightarrow & \frac{nl(l+1)}{2}V_{oj} - V_{il}V_{oj} > \frac{nj(j+1)}{2}V_{il} - V_{oj}V_{il} \\ \Leftrightarrow & \frac{l(l+1)}{2}V_{oj} > \frac{j(j+1)}{2}V_{il} \Leftrightarrow \frac{V_{oj}}{\frac{j(j+1)}{2}} > \frac{V_{il}}{\frac{l(l+1)}{2}} \end{aligned}$$

for all $j, l \in \{1, \dots, k\}$, we have

$$\begin{aligned} & \max_{j=1, \dots, k} \frac{V_{oj}}{\frac{nj(j+1)}{2} - V_{oj}} > \max_{j=1, \dots, k} \frac{V_{ij}}{\frac{nj(j+1)}{2} - V_{ij}} \\ \Leftrightarrow & \max_{j=1, \dots, k} \frac{V_{oj}}{\frac{j(j+1)}{2}} > \max_{j=1, \dots, k} \frac{V_{il}}{\frac{l(l+1)}{2}}; \end{aligned}$$

that is,

$$\widehat{Z}^*_o > \widehat{Z}^*_i \Leftrightarrow Z_o > Z_i. \quad \square$$

Proof of Theorem 2.

$$\begin{aligned}
\bar{Z}_o &= \frac{1}{\varepsilon_o^*} \int_0^{\varepsilon_o^*} \widehat{Z}_o^*(\varepsilon) d\varepsilon \\
&= \frac{1}{\varepsilon_o^*} \int_0^{\varepsilon_o^*} \left((m-1)V_o^* + \varepsilon \left(V_{ok} - V_o^* \left(\frac{nk(k+1)}{2} - V_{ok} \right) \right) \right) d\varepsilon \\
&= (m-1)V_o^* + \left(V_{ok} - V_o^* \left(\frac{nk(k+1)}{2} - V_{ok} \right) \right) \frac{1}{\varepsilon_o^*} \int_0^{\varepsilon_o^*} \varepsilon d\varepsilon \\
&= (m-1)V_o^* + \left(V_{ok} - V_o^* \left(\frac{nk(k+1)}{2} - V_{ok} \right) \right) \frac{\varepsilon_o^*}{2} \\
&= (m-1)V_o^* + \left(V_{ok} - V_o^* \left(\frac{nk(k+1)}{2} - V_{ok} \right) \right) \frac{m-1}{2 \left(\frac{nk(k+1)}{2} - V_{ok} \right)} \\
&= (m-1)V_o^* + \frac{m-1}{2} \frac{V_{ok}}{2 \left(\frac{nk(k+1)}{2} - V_{ok} \right)} - \frac{m-1}{2} V_o^* \\
&= \frac{m-1}{2} \left(V_o^* + \frac{V_{ok}}{\frac{nk(k+1)}{2} - V_{ok}} \right) = (m-1) \frac{V_o^* + V_o^k}{2}. \quad \square
\end{aligned}$$

References

- Amin GR, Sadeghi H (2010) Application of prioritized aggregation operators in preference voting. *Int J Intell Syst* 25(10):1027–1034
- Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. *Manag Sci* 39(10):1261–1264
- Contreras I (2011) A DEA-inspired procedure for the aggregation of preferences. *Expert Syst Appl* 38(1):564–570
- Contreras I, Hinojosa MA, Mármol AM (2005) A class of flexible weight indices for ranking alternatives. *IMA J Manag Math* 16(1):71–85
- Cook WD, Kress M (1990) A data envelopment model for aggregating preference rankings. *Manag Sci* 36(11):1302–1310
- Ebrahimnejad A (2012) A new approach for ranking of candidates in voting systems. *OPSEARCH* 49(2):103–115
- Fishburn PC (1974) Paradoxes of voting. *Am Polit Sci Rev* 68(2):537–546
- Foroughi AA, Aouni B (2012) New approaches for determining a common set of weights for a voting system. *Int Trans Oper Res* 19(4):521–530
- Foroughi AA, Tamiz M (2005) An effective total ranking model for a ranked voting system. *Omega* 33(6):491–496
- Foroughi AA, Jones DF, Tamiz M (2005) A selection method for a preferential election. *Appl Math Comput* 163(1):107–116
- Green RH, Doyle JR, Cook WD (1996) Preference voting and project ranking using DEA and cross-evaluation. *Eur J Oper Res* 90(3):461–472
- Hadi-Vencheh A (2014) Two effective total ranking models for preference voting and aggregation. *Math Sci* 8(1):1–4, Article 115
- Hashimoto A (1997) A ranked voting system using a DEA/AR exclusion model: A note. *Eur J Oper Res* 97(3):600–604

- Hashimoto A, Wu DA (2004) A DEA-compromise programming model for comprehensive ranking. *J Oper Res Soc Jpn* 47(2):73–81
- Hosseinzadeh Lotfi F, Fallahnejad R (2011) A note on “A solution method to the problem proposed by Wang in voting systems”. *Appl Math Sci* 5(62):3051–3055
- Hosseinzadeh Lotfi F, Rostamy-Malkhalifeh M, Aghayi N, Ghelej Beigi Z, Gholami K (2013) An improved method for ranking alternatives in multiple criteria decision analysis. *Appl Math Modell* 37(1-2):25–33
- Llamazares B, Peña T (2009) Preference aggregation and DEA: An analysis of the methods proposed to discriminate efficient candidates. *Eur J Oper Res* 197(2):714–721
- Llamazares B, Peña T (2013) Aggregating preferences rankings with variable weights. *Eur J Oper Res* 230(2):348–355
- Llamazares B, Peña T (2015a) Positional voting systems generated by cumulative standings functions. *Group Decis Negot* 24(5):777–801
- Llamazares B, Peña T (2015b) Scoring rules and social choice properties: some characterizations. *Theory Decis* 78(3):429–450
- Noguchi H, Ogawa M, Ishii H (2002) The appropriate total ranking method using DEA for multiple categorized purposes. *J Comput Appl Math* 146(1):155–166
- Obata T, Ishii H (2003) A method for discriminating efficient candidates with ranked voting data. *Eur J Oper Res* 151(1):233–237
- Soltanifar M, Ebrahimnejad A, Farrokhi MM (2010) Ranking of different ranking models using a voting model and its application in determining efficient candidates. *Int J Soc Syst Sci* 2(4):375–389
- Stein WE, Mizzi PJ, Pfaffenberger RC (1994) A stochastic dominance analysis of ranked voting systems with scoring. *Eur J Oper Res* 74(1):78–85
- Wang NS, Yi RH, Liu D (2008) A solution method to the problem proposed by Wang in voting systems. *J Comput Appl Math* 221(1):106–113
- Wang YM, Chin KS (2007) Discriminating DEA efficient candidates by considering their least relative total scores. *J Comput Appl Math* 206(1):209–215
- Wang YM, Chin KS, Yang JB (2007a) Three new models for preference voting and aggregation. *J Oper Res Soc* 58(10):1389–1393
- Wang YM, Luo Y, Hua Z (2007b) Aggregating preference rankings using OWA operator weights. *Inform Sci* 177(16):3356–3363
- Wu J, Liang L, Zha Y (2009) Preference voting and ranking using DEA game cross efficiency model. *J Oper Res Soc Jpn* 52(2):105–111