A NOTE ON COMPOSED SCORING RULES

HATEM SMAOUI

CERESUR, Université de La Réunion

Abstract. Composed scoring rules are related to important characterization results in social choice theory (Smith, 1973; Young, 1974, 1975). This shows the theoretical interest of this class of aggregation methods. At a first glance, they appear to be also of practical interest as the use of several vectors minimizes the number of ties in the social preference ordering. We show that it is not necessary, in real situations, to resort to composed scoring rules to achieve this goal: whenever a finite upper bound is placed on the number of voters (which is always possible in practice), composed scoring rules revert to simple scoring rules.

keywords: Voting theory, scoring rules, consistency.

JEL Classification: D71, D72.

1. Introduction

One of the important results of social choice theory is the axiomatic derivation of the class of (simple or composed) scoring rules, provided by Smith (1973) and Young (1974, 1975). Under a simple scoring rule, the candidates receive points based on their position on each individual preference according to a pre-fixed scoring vector. Then they are ranked by their total number of points. Composed scoring rules, which allow a hierarchy of tie-breaking scoring vectors, are generally considered as a "refinement" of simple scoring rules: ties generated by the first vector are decided according to the second vector, and so forth.

In their characterization theorems, Smith and Young assume that the number of voters is unbounded: the voting procedures are asked to work for all possible finite sizes of the electorate. This essentially permits to require criteria such as separability (which is the principal property of scoring rules) and continuity (which distinguishes between simple and composed scoring rules). However, in real-world elections, it is always possible to suppose that the number of voters does not exceed a certain finite upper bound. The object of this note is to prove a practical result showing that when such an hypothesis is made, any composed scoring rule become

Date: December 11, 2008.

We thank Annick Laruelle for her assistance. Helpful comments by Dominique Lepelley are gratefully acknowledged.

equivalent to (and thus can be replaced with) a simple scoring rule[1]. We give a constructive proof describing how to replace a sequence of scoring vectors by an appropriate single scoring vector.

2. **Scoring rules with unbounded number of voters**

We consider anonymous voting procedures on a set X of m candidates ($m \geq 2$). *Anonymity* means that names of voters should not matter in the collective decision (the election outcome is immune to permutations of individuals). Each voter ranks the m alternatives from her first choice to her least preferred candidate (no indifference is allowed). Let $P_1, \ldots, P_{m!}$ be an enumeration of the $m!$ linear orders (transitive, complete and anti symmetric binary relations) on X. For a finite electorate (or population) formed by n voters, the individual preferences can be collected in an (anonymous) profile described by an $m!$-tuple $\pi = (n_i)_{1 \leq i \leq m!}$, where n_i is the number of voters expressing a preference for ranking P_i and $\sum_{i=1}^{m!} n_i = n$. We denote by $\mathcal{D}$ the set of all possible profiles for all possible sizes of the population. The sum of two profiles, π and π', describing the preferences of two disjoint electorates is the profile $\pi + \pi'$ defined by the sum of the two $m!$-tuples representing π and π'. Let $\mathcal{R}(X)$ be the set of all weak orders (transitive, complete and reflexive binary relations) on X. An (anonymous) social welfare function (SWF) is a mapping that assigns to each profile π in $\mathcal{D}$ a collective preference $f(\pi)$ in $\mathcal{R}(X)$. For a SWF f, alternatives x, y in X, and a profile $\pi \in \mathcal{D}$, $xf(\pi)y$ means that x is socially judged at least as good as y. We write $xP(f(\pi))y$ when x is strictly preferred to y, and $xI(f(\pi))y$ when x is indifferent to y.

We focus on the particular class of SWF that are usually referred to as *simple scoring rules* (SSR). A scoring vector for m-candidate elections is a vector $v = (v_k)_{1 \leq\ k \leq m}$ in $\mathbb{R}^m$: for each voter, an alternative receives v_k points if it is ranked kth by the voter[2]. The score $S_v(\pi, x)$ obtained by an alternative x for a profile π, using vector v is computed by the formula: $S_v(\pi, x) = \sum_{k=1}^{m} c_k(\pi, x) v_k$, where $c_k(\pi, x)$ shows how many times the alternative x is ranked kth in profile π. The simple scoring rule associated with the vector v is the SWF f_v defined by[3]:

$$\forall x, y \in X, \forall \pi \in \mathcal{D}, x f_v(\pi) y \Leftrightarrow S_v(\pi, x) \geq S_v(\pi, y).$$

The most used SSRs are the Borda count ($v = (m-1, m-2, \ldots, 1, 0)$), the plurality rule ($v = (1, 0, \ldots, 0)$) and the antiplurality rule ($v = (1, ..., 1, 0)$). It is easy to see

[1]This result was initially conjectured by William S Zwicker (private communication, 2006).

[2]We do not require scoring vectors to be monotonic (i.e. the points for first place are at least as large as those for second place, and so forth).

[3]Simple scoring rules can also be defined in terms of social choice correspondences (SCC) which are functions that associate a non empty subset of candidates to each possible profile. In this case, a SSR is defined as a SCC applying a scoring vector and choosing, for each profile, all candidates with the highest score. The result established in this note also holds when SSRs are defined in the context of SCCs.

that $f_v = f_w$ whenever w is a positive linear transformation of v ($w = av + be$, where $a, b \in \mathbb{R}, a > 0$ and e is the vector of $\mathbb{R}^m$ defined by $e = (1, \ldots, 1)$).

Composed scoring rules[4] (CSR) are a refinement of SSRs, in the sense that they permit to reduce the number of ties in the social outcome. They were introduced by Smith (1973) and by Young (1974, 1975). A composed scoring rule works as follows: a first scoring vector v^1 is used to rank the m candidates. Ties produced by f_{v^1} are resolved by the mean of a second vector v^2 (which is not a positive linear transformation of v^1). If there are still ties, a third scoring vector is used, and so forth.

Definition 1. *Let $(v^1, \cdots, v^t)$ be a sequence of t scoring vectors in $\mathbb{R}^m$ and let $f_{v^1}, \cdots, f_{v^t}$ be the associated simple scoring rules. The composed scoring rule associated with the sequence $(v^1, \cdots, v^t)$, denoted by $f_{v^t} \circ \cdots \circ f_{v^1}$, is the SWF f defined by: $\forall \pi \in \mathcal{D}, \forall x, y \in X$,*

- $xP(f(\pi))y \Leftrightarrow (\exists j \in \{1, \cdots, t\} : xI(f_{v^i}(\pi))y, \forall i < j$ *and* $xP(f_{v^j}(\pi))y)$,
- $xI(f(\pi))y \Leftrightarrow (xI(f_{v^i}(\pi))y, \forall i \in \{1, \cdots, t\})$.

According to Definition 1, simple scoring rules are particular cases of composed scoring rules when a single vector is used. Note that the composition of simple (or composed) scoring rules is "associative". This means that $f_{v^t} \circ \cdots \circ f_{v^1}$ can be obtained indistinctly as the composition of CSR $f_{v^t} \circ \cdots \circ f_{v^2}$ and SSR f_{v^1} (applying f_{v^1} at first), or as the composition of the two CSRs $f_{v^t} \circ \cdots \circ f_{v^3}$ and $f_{v^2} \circ f_{v^1}$ (applying $f_{v^2} \circ f_{v^1}$ at first), etc.

By definition, simple and composed scoring rules are anonymous. One can easily see that they also satisfy the *neutrality* condition, in the sense that candidates are treated equally. More precisely, if all voters permute the alternatives in the same way in their preferences, then the same permutation is performed in the outcome. The principal condition satisfied by SSRs and CSRs is the *separability* (or consistency) axiom: if alternative x is socially considered at least as good as alternative y in two disjoint electorates, this conclusion should remain true when the two populations are assembled. If, in addition, candidate x is strictly preferred to y by one of the two groups, the same collective result must hold for the whole population. More formally, a SWF f is consistent if, for all π, π' in $\mathcal{D}$ and for all x, y in X,

- $(xf(\pi)y$ and $xf(\pi')y) \Rightarrow xf(\pi + \pi')y$,
- $(xf(\pi)y$ and $xP(f(\pi'))y) \Rightarrow xP(f(\pi + \pi'))y$.

Smith (1973) is the first to characterize scoring rules, followed by Young (1975). In order to distinguish SSRs from CSRs, Smith introduced a condition called *Archimedean property* (continuity in Young's terminology). This condition states Roughly that a sufficiently large group of voters can impose its will on any other fixed-size group. Formally, a SWF f is Archimedean if: $\forall, \pi, \pi' \in \mathcal{D}$, $\forall, x, y \in X$,

[4] "Generalized point systems" in Smith's terminology and "scoring functions" in Young's terminology.

$$xP(f(\pi))y \Rightarrow (\exists\ q_0 \in \mathbb{N} : q > q_0 \Rightarrow xP(f(q\pi + \pi'))y).$$

Theorem 1 (Smith, 1973)**.** *A neutral and anonymous social welfare fonction is separable if only if it is a (simple or composed) scoring rule. It is also Archimedean if only if it is a simple scoring rule.*

3. **Bounded number of voters**

Theorem 1 (see also Young, 1974, 1975) shows the theoretical interest of composed scoring rules. At a first glance, they appear to be also of practical interest as the use of several vectors minimizes the number of ties in the social preference ordering. In fact, it is not necessary, in real situations, to resort to composed scoring rules to achieve this goal. Indeed, we will show that whenever a finite upper bound is placed on the number of voters (which is always possible in practice), composed scoring rules revert to simple scoring rules. In other words, all collective results produced by a CSR (applied to profiles not exceeding a certain size) can be obtained by the mean of a suitable SSR.

For an integer n ($n \geq 2$), let $\mathcal{D}_n$ be the set of all profiles on X of size less than or equal to n. Thus $\mathcal{D}_n$ is the set of all profiles $\pi = (n_i)_{1 \leq i \leq m!}$ with $\sum_{i=1}^{m!} n_i \leq n$.

Theorem 2. *When the domain of social welfare functions is restricted to $\mathcal{D}_n$, any composed scoring rule is equivalent to a simple scoring rule.*

For a SWF f, we denote by $f_{|\mathcal{D}_n}$ the restriction of f to $\mathcal{D}_n$. The proof of Theorem 2 is based on the following lemma.

Lemma 1. *Let $f = f_{v^2} \circ f_{v^1}$ be a CSR defined by the composition of two SSRs. For each integer n ($n \geq 2$), there exists a scoring vector v such that $f_{|\mathcal{D}_n} = f_{v|\mathcal{D}_n}$.*

Proof. If v^1 or v^2 is a vector with equal components, f trivially reverts to a simple scoring rule. Assume that v^1 and v^2 are not equal-components vectors. By invariance to positive affine transformations, we can suppose that all components of v^1 and v^2 are positives. Indeed, it is always possible to add a vector of the form $a(1, 1, \cdots, 1)$ with $a > 0$, to obtain positive components.
Since X and $\mathcal{D}_n$ are finite sets, we can define:

$$\mu = max\{S_{v^2}(\pi, x) : x \in X, \pi \in \mathcal{D}_n\}, \text{ and}$$
$$\epsilon = min\{S_{v^1}(\pi, x) - S_{v^1}(\pi, y) : x, y \in X, x \neq y, \pi \in \mathcal{D}_n \text{ and } S_{v^1}(\pi, x) > S_{v^1}(\pi, y)\}.$$

Notice that $\mu > 0$ and $\epsilon > 0$ (since v^1 and v^2 are not equal-components vectors). In fact, μ and ϵ can be easily calculated (For instance, $\mu = n \cdot \alpha$ where α is the greatest component of v^2). Let v be the vector defined by:

$$v = v^1 + \frac{\epsilon}{2\mu} v^2$$

We will show that the restrictions of f and f_v to $\mathcal{D}_n$ are equal. We begin by establishing that, for all x, y in X $(x \neq y)$ and for all π in $\mathcal{D}_n$,

$$S_{v^1}(\pi, x) > S_{v^1}(\pi, y) \Rightarrow S_v(\pi, x) > S_v(\pi, y) \qquad (1)$$

$$S_v(\pi, x) > S_v(\pi, y) \Rightarrow S_{v^1}(\pi, x) \geq S_{v^1}(\pi, y) \qquad (2)$$

$$S_v(\pi, x) = S_v(\pi, y) \Rightarrow S_{v^1}(\pi, x) = S_{v^1}(\pi, y) \qquad (3)$$

$$S_v(\pi, x) = S_v(\pi, y) \Rightarrow S_{v^2}(\pi, x) = S_{v^2}(\pi, y) \qquad (4)$$

(1). If $S_{v^1}(\pi, x) > S_{v^1}(\pi, y)$, then

$$\begin{aligned} S_v(\pi, x) - S_v(\pi, y) &= (S_{v^1}(\pi, x) - S_{v^1}(\pi, y)) + \frac{\epsilon}{2\mu}(S_{v^2}(\pi, x) - S_{v^2}(\pi, y)) \\ &> \qquad \frac{\epsilon}{2} \qquad + \frac{\epsilon}{2\mu}(-\mu) \qquad (= 0) \end{aligned}$$

(2). If $S_v(\pi, x) > S_v(\pi, y)$, suppose $S_{v^1}(\pi, x) < S_{v^1}(\pi, y)$. Applying (1), we obtain $S_v(\pi, x) < S_v(\pi, y)$. Contradiction.
(3). It is also a consequence of (1): if $S_v(\pi, x) = S_v(\pi, y)$, suppose for example $S_{v^1}(\pi, x) > S_{v^1}(\pi, y)$. Again, by applying (1) we obtain $S_v(\pi, x) > S_v(\pi, y)$. Contradiction.
(4). If $S_v(\pi, x) = S_v(\pi, y)$, by (3) we obtain $S_{v^1}(\pi, x) = S_{v^1}(\pi, y)$. It follows from the definition of vector v that $S_{v^2}(\pi, x) = S_{v^2}(\pi, y)$.

We can now show that f and f_v coincide on $\mathcal{D}_n$. We have to establish that for all $x, y \in X$ $(x \neq y)$ and for all $\pi \in \mathcal{D}_n$,

$$xP(f(\pi))y \Rightarrow xP(f_v(\pi))y \qquad (a)$$

$$xI(f(\pi))y \Rightarrow xI(f_v(\pi))y \qquad (b)$$

$$xP(f_v(\pi))y \Rightarrow xP(f(\pi))y \qquad (c)$$

$$xI(f_v(\pi))y \Rightarrow xI(f(\pi))y \qquad (d)$$

(a). If $xP(f(\pi))y$, by the definition of f, we have:
(i) $S_{v^1}(\pi, x) > S_{v^1}(\pi, y)$ or
(ii) $S_{v^1}(\pi, x) = S_{v^1}(\pi, y)$ and $S_{v^2}(\pi, x) > S_{v^2}(\pi, y)$.
In case (i), (1) gives $S_v(\pi, x) > S_v(\pi, y)$. In case (ii), by definition of v, we obtain $S_v(\pi, x) > S_v(\pi, y)$. In both cases, we have $xP(f_v(\pi))y$.
(b). If $xI(f(\pi))y$, we have: $S_{v^1}(\pi, x) = S_{v^1}(\pi, y)$ and $S_{v^2}(\pi, x) = S_{v^2}(\pi, y)$. Hence $S_v(\pi, x) = S_v(\pi, y)$ and thus $xI(f_v(\pi))y$.
(c). If $xP(f_v(\pi))y$, we have $S_v(\pi, x) > S_v(\pi, y)$.
By (2), we have $S_{v^1}(\pi, x) \geq S_{v^1}(\pi, y)$. If $S_{v^1}(\pi, x) > S_{v^1}(\pi, y)$ then $xP(f(\pi))y$. If $S_{v^1}(\pi, x) = S_{v^1}(\pi, y)$, by definition of v, we must have $S_{v^2}(\pi, x) > S_{v^2}(\pi, y)$, thus $xP(f(\pi))y$.
(d). If $xI(f_v(\pi))y$, we have $S_v(\pi, x) = S_v(\pi, y)$. By (3) and (4), we obtain $xI(f(\pi))y$. □

Proof of Theorem 2. Let $f = f_{v^t} \circ \cdots \circ f_{v^1}$ be a composed scoring rule. For $t = 2$, the result is established in Lemma 1. Suppose $t > 2$. By Lemma 1 we can replace, on $\mathcal{D}_n$, $f_{v^2} \circ f_{v^1}$ with f_v, for some scoring vector v. Thus the expression of f reduces to $f = f_{w^t} \circ \cdots \circ f_v$. By applying recursively Lemma 1, f completely reduces to a simple scoring rule. □

Example 1. (1) *In a three-candidate election, let f be the CSR defined by $f = f_{v^2} \circ f_{v^1}$, where $v^1 = (2, 1, 0)$ and $v^2 = (1, 0, 0)$: candidates are ranked by their Borda score and possible ties are resolved according to plurality rule. If the maximal size of the electorate is* 100*, constants μ and ϵ defined in the proof of Lemma 1 are given by $\mu = 100$ and $\epsilon = 1$. Thus, f can be replaced with the SSR associated with the scoring vector $v = (2 + \frac{1}{200}, 1, 0)$.*

(2) *Consider a four-candidate election with at most* 1000 *voters. Let g be the CSR defined by $g = f_{v^3} \circ f_{v^2} \circ f_{v^1}$ where $v^1 = (3, 2, 1, 0)$, $v^2 = (1, 0, 0, 0)$ and $v^3 = (0, 1, 0, 0)$. The CSR $f_{v^2} \circ f_{v^1}$ can be replaced with the SSR using vector $v = (3 + \frac{1}{2000}, 2, 1, 0)$ ($\mu = 1000$ and $\epsilon = 1$). Hence g is equivalent to $f_{v^3} \circ f_v$ and finally reverts to the SSR using vector $w = (3 + \frac{1}{2000}, 2 + \frac{1}{4000000}, 1, 0)$ ($\mu' = 1000$, $\epsilon' = \frac{1}{2000}$).*

4. Concluding remark

Our result is obtained by assuming the realistic hypothesis of a bounded number of voters. Under this assumption, there is no difference between simple and composed scoring rules and the Archimedean property used in Theorem 1 becomes superfluous. An interesting (but probably difficult) problem is to adapt the definition of separability to the context of bounded-size electorate and to reformulate Smith and Young characterization of scoring rules.

References

Smith, J.H. (1973) "Aggregation of preferences with variable electorate", *Econometrica* **41**, 1027-1041.

Young, H.P. (1974) "A Note on Preference Aggregation", *Econometrica* **42**, 1129-1131.

Young, H.P. (1975) "Social choice scoring functions", *SIAM Journal of Applied Mathematics* **28**, 824-838.