CENTRE DE RECHERCHE SUR LES STRATÉGIES ÉCONOMIQUES

Combining diversity and excellence in multiwinner elections

Mostapha Diss, Clinton Gubong Gassi, Issofa Moyouwou

September 2023

Working paper No. 2023 – 05

CRESE
30, avenue de l'Observatoire
25009 Besançon
France
http://crese.univ-fcomte.fr/

UFR SJEPG
Sciences juridiques économiques politiques et de gestion
UNIVERSITé DE FRANCHE-COMTé

Combining diversity and excellence in multiwinner elections*

Mostapha Diss[†‡] Clinton Gubong Gassi[§¶] Issofa Moyouwou[‖]

September 26, 2023

Abstract

We address the problem of electing a committee subject to diversity constraints. Given a set of candidates and a set of voters, such that each voter is represented by a linear order, the goal is to select a fixed-size subset of candidates by combining the excellence of candidates and a given form of diversity requirements. The grounding assumption in this paper is that the set of candidates is slotted into at least two groups according to a specific attribute such as gender, religion, ethnicity, or profession, and the diversity constraint takes the form of a vector of integers specifying the lowest number of candidates required from each group. We introduce a class of voting rules suitable for electing a diverse committee in this framework and we show how this class of rules handles the issue of combining both excellence and diversity. Furthermore, we provide some axiomatic properties that highlight the behavior of these rules when we aim to select a diverse committee.

Keywords: Voting, multiwinner elections, committee, diversity, axioms.
JEL classification: D71, D72.

*We would like to thank Sylvain Béal, Egor Ianovski, Eric Kamwa, Aleksei Kondratev, and Abdelmonaim Tlidi for their valuable comments and useful advice on our manuscript. This article has been presented in various conferences and workshops: CEPR Workshop on Democracy, ETH Zurich, Switzerland; MDOD, University of Paris 1 Panthéon-Sorbonne; CRESE, University of Franche-Comté, Besançon, France; University Mohammed VI Polytechnic, Rabat, Morocco. We thank all the participants for their comments. Mostapha Diss would like to acknowledge financial support from Région Bourgogne Franche-Comté within the program ANER 2021-2024 (project DSG).

†Université de Franche-Comté, CRESE, F-25000 Besançon, France. Email: mostapha.diss@univ-fcomte.fr.

‡Africa Institute for Research in Economics and Social Sciences (AIRESS), University Mohamed VI Polytechnic, Rabat, Morocco.

§Université de Franche-Comté, CRESE, F-25000 Besançon, France. Email: clinton.gassi@univ-fcomte.fr.

¶Département de Mathématiques - Faculté des sciences - Université of Yaoundé I. BP 812 Yaoundé, Cameroon.

‖Ecole Normale Supérieure - Département de Mathematiques - Université de Yaoundé I. BP 47 Yaoundé, Cameroon. Email: issofa.moyouwou@ens.cm.

1 Introduction

The problem of selecting a group of items or individuals from a larger set is an ubiquitous problem in both theory and real life. Such a problem can arise in parliamentary elections, shortlisting candidates for a competition, choosing a set of movies to be offered on airplanes, selecting a number of items to be put on a page of an online store, etc. In social choice theory, we generally talk about a set of voters (or individuals) who express their preferences over a set of candidates (or alternatives) in order to select a fixed-size subset of candidates called a *committee*. These kinds of problems are well known under the name of *multiwinner elections* or *committee selection*.

There are many multiwinner voting rules based on many different ideas and principles. The more studied ones are those based on k-winner extensions of well-known single-winner rules, where k is the target size of the committee to be selected. In this category, the family of rules that has received the most interest in the literature is undoubtedly the family of *committee scoring rules* that we will focus on in the paper at hand. This family of rules was introduced by Elkind et al. (2017) as the multiwinner analogues of the well-known single-winner positional scoring rules. The idea of single-winner positional scoring rules (or, simply, *scoring rules*) is to assume that each voter ranks all the candidates from the most to the least preferred one and to consider a scoring vector that associates each position in a ranking with a given score value; the final score of a candidate is calculated as the sum of the scores he/she (hereinafter she) obtains from all the voters; finally, the candidate with the highest total score wins the election.[1] Roughly speaking, and similarly to the single-winner setting, under a committee scoring rule each voter assigns a predefined score to each committee based on the positions of the committee members in the considered voter's ranking and, at the end, the winning committee is the one with the maximum total score computed as the sum of the scores received from all voters to every candidate. A formal definition of committee scoring rules will be given later. Note that Skowron et al. (2019) provided an axiomatic characterization of committee scoring rules which extends by the way the characterizations given by Young (1974, 1975) for single-winner scoring rules. Due to their simplicity, the most commonly used committee scoring rules are undoubtedly those that are (*weakly*) *separable*. Under (weakly) separable committee scoring rules, candidates are one by one rated according to a single-winner scoring vector and any set of k candidates with the highest total scores is picked at the end as a winning committee. We refer the reader to Elkind et al. (2017) and Faliszewski et al. (2018, 2019), among others, for further discussions about committee scoring rules in general and (weakly) separable committee scoring rules in particular.

[1] With a possible tie-breaking rule which could be used to break ties among winning candidates having the same highest total score.

Multiwinner elections with diversity constraints

Choosing a committee can sometimes be subject to certain constraints based on different ideas and goals that should be achieved as argued by Brams (1990), Kamada and Kojima (2015), Lu and Boutilier (2011), and Skowron et al. (2016), among others. The paper at hand is concerned with the question of diversity. Precisely, the general idea of this piece of literature is that there is one (or more) attribute(s), which can be gender, age, ethnicity, etc., allowing the set of candidates to be sorted into several classes (or types), such that each candidate is labeled by one (or more) type(s) according to each attribute.[2] Then, the goal is to ensure a certain degree of diversity among the candidates composing the selected committee. There are numerous examples of such constraints. For instance, shortlisting candidates for a competition can be subject to a condition requiring an equal number of men and women, or a condition seeking to achieve a certain regional equilibrium if the candidates' home region is an important aspect of the committee selection so that no region feels excluded or underrepresented. Another example has to do with the choice of movies to be offered on airplanes; since the goal is to please all the passengers, it would be "fair" that plane movies include thrillers, biographical dramas, romantic comedies or other varieties. In social choice theory, we generally say that the committee to be selected is subject to a diversity constraint and we talk about the selection of a *diverse committee*.

The literature devoted to the framework of diverse committee selection is recent and rich, and to the best of our knowledge it may be classified into two groups.

The first group concerns the multi-attribute framework, assuming that each candidate is represented by one label over each attribute. In this framework, Bredereck et al. (2018) for instance considered that the diversity constraint is expressed by two vectors of integers specifying the lowest and the highest numbers of candidates to be selected from each attribute class and the goal is to optimize a given objective function so as to provide a committee that meets in the best possible way the diversity constraint. The authors mainly focused on the computational complexity of selecting such a committee. In the same vein, Lang and Skowron (2018) studied the problem of selecting a diverse committee by assuming that the set of candidates is slotted into many classes according to many attributes. The diversity constraint is considered in their model as a sequence of vectors expressing the number of candidates to be selected from each class and each attribute, and the ultimate goal is to select a committee that is closer to the target sequence of vectors using some specific metrics. The authors also studied the computational complexity of managing such a problem, and furthermore provided some properties that should be satisfied when selecting a diverse committee in this framework. Many other works have addressed the same problem in the multi-attribute setting, we refer the reader to Bei et al. (2022), Celis et al. (2017), and Do et al. (2021), among others.

The second group consists of the setting where the set of candidates is slotted into several classes according to a single specific attribute. In this framework, Ianovski (2022)

[2]Obviously, the classes can overlap or not depending on the considered attribute.

for instance considered two types of diversity constraints, namely the *interval constraints* and *dominance constraints*,[3] and studied the problem of selecting a feasible committee that preserves the excellence of candidates under an objective function. Ianovski (2022) then focused on the computational complexity of this problem. A second example concerns the work of Aziz (2019) who considered that the set of candidates is structured into several non-disjoint classes according to a specific attribute, and defined a diversity constraint as a vector of integers specifying the lowest number of candidates to be selected from each class. The author assumed that the preferences of voters are already aggregated into a single (weak) order on the set of candidates and provided an algorithm that combines both excellence (measured according to the positions of the candidates in the social weak order) and diversity. Excellence and diversity are modelled by two axioms, namely *justified envy-freeness* and *type optimality* that we will present and discuss in detail later in this paper. Many other works are dedicated to the same framework; we can point the reader toward the works of Kagita et al. (2021), Relia (2021), and Thejaswi et al. (2021), among others.

Our contribution

We consider the framework where (i) the set of candidates is partitioned into at least two disjoint groups according to a single specific attribute, and (ii) the diversity constraint takes the form of a vector of integers specifying the lowest number of candidates that are required from each group. Our main objective is to select a committee which effectively combines both the excellence of candidates in terms of scores and the diversity requirements. We consider the well-known committee scoring rules and we introduce a diversity reward that we distribute to each committee according to its degree of diversity, i.e., its ability to meet the diversity constraint. This allows us to define a new class of committee selection rules that we consider as suitable for selecting a diverse committee in this framework. We call it the class of *diverse committee scoring rules*. We further provide some desirable properties that are satisfied by this class of rules. Some of these properties are adjustments of well-known axioms already considered in the literature, mainly by Aziz (2019) and Elkind et al. (2017), whereas other axioms are proper to our model. We pay a particular attention to the diverse committee scoring rules that are based on (weakly) separable scoring functions and we characterize this subclass of rules using a particular property.

The paper is organized as follows: Section 2 presents the model by laying out some basic definitions. Section 3 is devoted to the definition of the class of diverse committee scoring rules and describes how it operates. In Section 4, we provide some axioms that we see as suitable when the goal is to select a diverse committee and we evaluate the class of rules that we propose in the light of those axioms. Section 5 concludes and presents some directions for future research.

[3]The interval constraints indicate the interval to which the number of candidates to be selected from each class must belong. The dominance constraint shows, for any two classes, which one deserves more candidates to be selected.

2 Preliminary definitions

Consider a non-empty set A of m alternatives (or candidates) and a non-empty set N of n voters (or individuals) with $m \geq 3$ and $n \geq 2$. Alternatives are denoted by small letters a, b, c, ..., or a_1, a_2, a_3, etc. Voters are denoted by positive integers 1, 2, 3, etc. We denote by $\mathbb{N}$ the set of all non-negative integers and by $\mathbb{N}^*$ the set of all positive integers. Throughout the paper, we simply write $[r]$ to denote the set $\{1, \ldots, r\}$ for any positive integer $r \in \mathbb{N}^*$. We assume that the set of alternatives is slotted into l subsets $A_1, \ldots, A_l$, with $l \geq 2$, according to a single specific attribute (e.g., gender, religion, political orientation, etc.) such that $A = \cup_{j=1}^{l} A_j$. For every type $j \in [l]$, we denote by $m_j = |A_j|$ the cardinality of the class A_j. Although some classes can overlap,[4] we focus in this paper on attributes for which any candidate belongs to only one class; that is, $A_j \cap A_{j'} = \emptyset$ for all $j, j' \in [l]$ such that $j \neq j'$. For any alternative $a \in A$, we denote by $j(a)$ the type of a; that is, the integer from $[l]$ such that $a \in A_{j(a)}$. We say that $A_{j(a)}$ is the class of candidate a. For every A_j, we use throughout the paper the term *type* referring to j and the term *class* referring to the subset A_j.

We assume that every voter ranks all the alternatives from the most preferred to the least preferred one without the possibility of ties. Every individual preference is then a linear order on A; that is, a complete, anti-symmetric, and transitive binary relation on A. The set of all linear orders on A is denoted by $\mathcal{P}$. It is worth mentioning that the partition of the set of candidates into classes does not affect the ranking of any voter; that is, each voter casts her ranking regardless of the types of the candidates. Given $i \in N$, the (linear) ranking or the preference relation of i is denoted by p_i. The n-tuple $p = (p_1, p_2, \ldots, p_n)$ which specifies the ranking of each voter is called a preference profile (or simply a profile). The set of all profiles with n voters is denoted by $\mathcal{P}^n$. For any two alternatives a and b, we write $a \succ_i b$ or simply ab (in row or column) if voter i strictly prefers a to b.

Definition 1 *Given a voter i, the rank $r(p_i, a)$ of any alternative $a \in A$ in the preference relation p_i is defined by*

$$r(p_i, a) = \Big| \{b \in A : b \succ_i a\} \Big| + 1 = m - \Big| \{b \in A : a \succ_i b\} \Big|. \tag{1}$$

Each individual preference p_i induces a restricted ranking p_i^j (also denoted by $\succ_i^j$) over each class A_j such that, for all $a, b \in A_j$,

$$a \succ_i^j b \Longleftrightarrow a \succ_i b. \tag{2}$$

For instance, assume that $A = \{a, b, c, d, e\}$, $l = 2$, $A_1 = \{a, c, e\}$, and $A_2 = \{b, d\}$. Let us consider the preference relation $p_i = adecb$ of a given voter i, meaning that the highest-ranked candidate is a, candidate d is ranked second, and so forth until the lowest-ranked

[4] The framework in which an individual can belong to two or more classes (e.g., spoken languages, citizenships, etc.) is out of the scope of this paper. See, for instance, Aziz (2019), Bei et al. (2022), Bredereck et al. (2018), Do et al. (2021), and Lang and Skowron (2018), among others.

candidate b. In this case, the preference relation p_i induces the two following restricted rankings over each class A_j for $j \in [2]$: $p_i^1 = aec$ and $p_i^2 = db$.

For any integer $k \in [m-1]$, we call a *committee* of size k any k-element subset of A and the set of all possible committees of size k for the set of candidates A is denoted by 2^A_k. Given a set of candidates, a set of voters, and a preference profile, we consider throughout this paper a framework in which the goal is to select a committee of size k taking into account both the excellence of candidates in terms of scores and the diversity of the selected committee with regards to the classes of candidates.

Definition 2 *We call a diversity constraint any quota vector $q = (q_1, \ldots, q_l) \in \mathbb{N}^l$ specifying the lowest number of candidates from each class that the selected committee should contain and we denote by $[l]^*_q$ the set $\{j \in [l] : q_j \neq 0\}$ of all types j for which the diversity constraint requires at least one candidate from A_j.*

For a given committee $W \in 2^A_k$ and a type $j \in [l]$, we denote by W^j the set of candidates from A_j belonging to W; that is, $W^j = W \cap A_j$. We say that a committee W satisfies the diversity constraint q if $|W^j| \geq q_j$ for all $j \in [l]$. Recalling that the committee size k is fixed and any two classes are disjoint, the condition $\sum_{j=1}^{l} q_j \leq k$ holds since the sum of the lowest number of candidates required from all the possible classes cannot exceed the target size of the committee. It is worth mentioning that there can be no committee satisfying the diversity constraint q, which means that imposing the diversity constraint makes the selection process unfeasible. This scenario happens when at least one class A_j does not contain enough candidates to achieve the enforced quota q_j for that class. In this case, the aim will be to select the committee that best enables us to get close to the diversity constraint. Thus, for any type $j \in [l]$, we denote $\delta(A_j) = \min\{|A_j|, q_j\}$ as the minimum between the lowest number q_j of candidates that is required from the class A_j and the total number of candidates available in A_j.

Definition 3 *Given a diversity constraint q, we say that a committee W is a q-diverse committee if $|W^j| \geq \delta(A_j)$ for all $j \in [l]$, and we denote by $2^A_{k,q}$ the set of all q-diverse committees of size k.*

Literally, $2^A_{k,q}$ stands for the set of committees of size k that are closest to the diversity constraint, and this set coincides with the set of size-k committees satisfying the diversity constraint whenever such committees exist.

Example 1 *Consider a set of candidates $A = \{a, b, c, d, e\}$ partitioned according to gender (for instance), with $A_1 = \{a, b, c\}$ being the set of all available men and $A_2 = \{d, e\}$ being the set of all available women. Suppose that $k = 3$ and $q = (1, 2)$ requiring exactly one man and two women. In this case, the diversity constraint can be satisfied and the set $2^A_{k,q}$ contains the committees $\{a, d, e\}$, $\{b, d, e\}$, and $\{c, d, e\}$. Assume now that $k = 4$ and $q = (1, 3)$ requiring exactly one man and three women. In this case, the diversity constraint cannot be satisfied and the set $2^A_{k,q}$ contains the three closest committees to the diversity constraint, which are $\{a, b, d, e\}$, $\{a, c, d, e\}$, and $\{b, c, d, e\}$.*

Definition 4 *A Diverse Committee Selection Rule (DCSR) is a mapping F that assigns to any profile p, any integer $k \in [m-1]$, and any diversity constraint q, one (or more) winning committee(s) of size k. The set of winning committees with respect to the triplet (p, k, q) is denoted by $F(p, k, q)$ and it is called the social outcome of the triplet (p, k, q) when the DCSR is F.*

3 Committee scoring rules and diversity constraints

Our main task in this section is to define a diverse form of committee scoring rules that we tentatively see as suitable for selecting a diverse committee in our framework. We start by recalling the definition of committee scoring rules given by Elkind et al. (2017).

3.1 Classical committee scoring rules

Given a committee $W \in 2^A_k$, the rank $r(p_i, W)$ of the committee W with regards to the linear order p_i of a given voter i is the increasing sequence $(i_1, \ldots, i_k)$ obtained by ordering the set $\{r(p_i, a) : a \in W\}$. For instance, assume that the set of candidates is $A = \{a, b, c, d, e\}$, the preference of voter i is $p_i = bcade$, and the committee size is $k = 3$. Then, the rank of the committee $W = \{a, c, e\}$ in voter's i preference relation is $r(p_i, W) = (2, 3, 5)$. We denote by $[m]_k$ the set of all possible increasing sequences of k elements from $[m]$. In other words, $[m]_k$ stands for the set of all possible ranks of a given size-k committee in a given linear ranking. Given two committee ranks $I = (i_1, \ldots, i_k)$ and $J = (j_1, \ldots, j_k)$, we say that I *dominates* J, which is denoted by $I \succeq J$, if $i_t \leq j_t$ for all $t \in [k]$. In particular, $I_0 = (1, \ldots, k)$ dominates any other rank, and $J_0 = (m-k+1, \ldots, m)$ is dominated by any other rank.

Definition 5 *A committee scoring function is a function $f_{m,k} : [m]_k \to \mathbb{R}_+$ such that for all $I, J \in [m]_k$, $I \succeq J$ implies $f_{m,k}(I) \geq f_{m,k}(J)$. Given a committee scoring function $f_{m,k}$, the score of a committee $W \in 2^A_k$ with respect to the committee scoring function $f_{m,k}$ and a preference profile p is defined by*

$$S_{f_{m,k}}(p, W) = \sum_{i \in N} f_{m,k}\Big(r(p_i, W)\Big). \tag{3}$$

We can immediately deduce that the score of a committee W with respect to the committee scoring function $f_{m,k}$ always satisfies

$$n f_{m,k}(J_0) \leq S_{f_{m,k}}(p, W) \leq n f_{m,k}(I_0). \tag{4}$$

Definition 6 *A committee selection rule F is a committee scoring rule if there is a family of scoring functions $f = (f_{m,k})_{k \leq m-1}$ such that for any size $k \leq m-1$ and any preference profile p, the set of winning committees with respect to (p, k) is the set of all committees* *of size k with the highest score; that is,*

$$F(p,k)=\Big\{W\in 2^A_k : S_{f_{m,k}}(p,W)\geq S_{f_{m,k}}(p,W') \text{ for all } W'\in 2^A_k\Big\}. \tag{5}$$

As noted in the introductory section, the most studied committee scoring rules are undoubtedly the (weakly) separable committee scoring rules that rate the candidates separately according to a single-winner scoring vector and selects the k candidates with the highest scores.[5] Let us recall that a scoring vector for single-winner elections is a vector $\alpha=(\alpha_1,\dots,\alpha_m)$ of real-number scoring weights satisfying both $\alpha_1\geq\alpha_2\geq\cdots\geq\alpha_m$ and $\alpha_1>\alpha_m$. In other words, each voter gives α_1 points to her most favoured candidate, α_2 points to her second-ranked candidate, and so forth until α_m points to her bottom-ranked candidate. The (weakly) separable committee scoring rule associated to the family of scoring vectors $(\alpha^k)_{k\leq m-1}$ selects, for each committee size k, the k candidates with the highest number of total points since the total points collected determine a complete ranking of the candidates with possible ties. If the scoring vector α^k does not depend on k; that is $\alpha^k=\alpha^{k'}=\alpha$ for all $k,k'\leq m-1$, then the rule is simply said to be separable; otherwise, the accuracy *weakly* is needed. The vectors $\alpha^P=(1,0,0,\dots,0)$, $\alpha^B=(m-1,m-2,\dots,2,1,0)$, and $\alpha^{AP}=(1,1,\dots,1,0)$ are the scoring vectors associated with the well-known separable k-Plurality rule (also called *Single Non Transferable Voting rule*, SNTV), k-Borda rule, and k-Antiplurality rule respectively, whereas the vector $\alpha^{k,Bl}=(\underbrace{1,\dots,1}_{k\text{ times}},\underbrace{0,\dots,0}_{m-k\text{ times}})$ is the scoring vector used for the well-known weakly separable Bloc rule.

Formally, the score gained by any candidate $a\in A$ across a preference profile $p\in\mathcal{P}^n$ when the scoring vector α^k is considered is given by

$$S_{\alpha^k}(p,a)=\sum_{i=1}^{n}\alpha^k_{r(p_i,a)}. \tag{6}$$

Definition 7 *For a given profile $p\in\mathcal{P}^n$, the (weakly) separable committee scoring rule associated with the family of scoring vectors $(\alpha^k)_{k\leq m-1}$ outputs, for any committee size k, the committee(s) $W\in 2^A_k$ that maximise(s) the score $S_{\alpha^k}(p,W)$ defined by*

$$S_{\alpha^k}(p,W)=\sum_{a\in W}S_{\alpha^k}(p,a)=\sum_{a\in W}\sum_{i=1}^{n}\alpha^k_{r(p_i,a)}. \tag{7}$$

Every (weakly) separable committee scoring rule defined through a family of scoring vectors $(\alpha^k)_{k\leq m-1}$ is then a committee scoring rule associated with the family of scoring

[5]The subclass of separable committee scoring rules can be seen as the intersection between the committee scoring rules and the *candidate-wise* procedures already defined by Kilgour and Marshall (2012). A procedure is described as a candidate-wise procedure if the score of a given committee is the sum of the scores of all the candidates belonging to that committee, each of them considered as a 1-size committee. Note that the candidate-wise procedures are analogous to the additive procedures studied in a more general context in Kilgour (2010).

functions $(f_{m,k})_{k \le m-1}$ defined by

$$f_{m,k}(i_1, \dots, i_k) = \sum_{t=1}^{k} \alpha_{i_t}. \tag{8}$$

Note that the wide class of committee scoring rules contains other interesting rules which are not (weakly) separable. We can think about a version of the Chamberlin-Courant rule (Chamberlin and Courant, 1983) defined with the scoring function

$$f_{m,k}(i_1, \dots, i_k) = m - i_1. \tag{9}$$

Under this rule we start by fixing a scoring vector. For instance, the version defined in (9) is based on the Borda scoring vector. Then, each voter's score for a given committee with k candidates is defined to be the Borda score that she assigns to her most preferred candidate in that committee; the goal is then to find a committee that maximizes the joint scores collected from all voters.

3.2 Diverse committee scoring rules

As noted before, our proposal is to define a diverse form of committee scoring rules with a view to selecting a committee that allows to combine both the excellence of candidates in terms of scores and the diversity of its members. The main idea is to combine the score of any committee defined in (3) with another score reflecting the degree of diversity.

Definition 8 *Given a non-negative real number M and a family of committee scoring functions $f = (f_{m,k})_{k \le m-1}$, the diverse committee scoring rule associated to the couple (M, f), is the DCSR that outputs for each triplet $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, the set $F(p, k, q)$ of all committees $W \in 2^A_k$ that maximize the diverse score $DS_{f_{m,k}}(p, q, W)$ defined by*

$$DS_{f_{m,k}}(p, q, W) = S_{f_{m,k}}(p, W) + \sum_{j=1}^{l} \delta^j(W) M, \tag{10}$$

where $\delta^j(W) = \min\left\{ \left|W^j\right|, q_j \right\}$ for all $j \in [l]$.

The first component of $DS_{f_{m,k}}(p, q, W)$ has been defined in (3) and will be called throughout the paper the *score of excellence*. In other words, it represents the score that W would have obtained without the diversity requirement. The second component of $DS_{f_{m,k}}(p, q, W)$ is the additional score that has to do with the degree of diversity of the considered committee. It will be called the *total diversity reward* that is given to the committee W based on its ability to meet the diversity constraint q.[6] The basic idea behind

[6] Izsak et al. (2018) had a quite similar idea (but for a different model) to rate the committees based on (weakly) separable committee scoring rules and combined with *synergy functions* that describe how the candidates from different classes react each other into a committee. The goal of that model is to select a subset of candidates with a certain degree of affinities.

the second component of $DS_{f_{m,k}}(p,q,W)$ is the following: going through all possible types $j \in [l]$, the committee W receives a diversity reward of M as many times as $\delta^j(W)$. Note that more details on the choice of the reward M will be given later in the article. The variable $\delta^j(W)$ can be seen as the number of candidates from A_j that allow W to be considered as satisfactory regarding the requirement of diversity for type j. Indeed, if $|W^j| > q_j$ for a given type $j \in [l]$, since there are more candidates of type j than those required from that type, the committee W can be seen as satisfactory regarding the requirement of diversity of type j, thanks to the required q_j candidates of W^j. Now, if $|W^j| \leq q_j$, the committee W can also claim to have a certain degree of satisfaction concerning the requirement of diversity of type j, thanks to the $|W^j|$ candidates of W^j. Roughly, for any committee $W \in 2^A_k$, there are $\delta^j(W)$ candidates that allow W to claim a certain satisfaction with regards to the diversity requirement for type j. The total diversity reward that will be given to W for type j is then proportional to $\delta^j(W)$. The second component of $DS_{f_{m,k}}(p,q,W)$ is the sum of the diversity rewards collected over all types. Note finally that any two committees having the same diversity structure earn the same total diversity reward.[7]

A natural question is how the choice of the diversity reward M impacts the composition of the winning committee(s) derived from (10). Of course, if we consider $M = 0$ (or very small values of M), we simply recover the classical committee scoring rules. The first result of the paper at hand shows that, once the diversity reward M has reached a certain value that depends on the scoring function in consideration, the election process can be reduced to the selection of the most excellent committee(s) from $2^A_{k,q}$; that is, the committee(s) which meet(s) or which are closest to the diversity constraint and having the highest excellence score.

Theorem 1 *Given a family $f = (f_{m,k})_{1\leq k\leq m-1}$ of scoring functions, there exists a diversity reward $M(f)$ such that for any diverse committee scoring rule F defined by f and a diversity reward M with $M > M(f)$, it holds that for any $W \in 2^A_k$ and any triplet $(p,k,q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, $W \in F(p,k,q)$ if and only if W is a q-diverse committee with the highest (excellence) score.*

Proof. Let F be a diverse committee scoring rule associated to the couple (M,f) such that the collection f of scoring functions and the diversity reward M satisfy $M > M(f)$ where

$$M(f) = n\Big(f_{m,k}(I_0) - f_{m,k}(J_0)\Big). \tag{11}$$

Let p be a profile, $k \in [m-1]$ be a given committee size, q be a diversity constraint and $W \in 2^A_k$. We have to prove that $W \in F(p,k,q)$ if and only if $W \in 2^A_{k,q}$ and $S_{f_{m,k}}(p,W) \geq S_{f_{m,k}}(p,W')$ for all $W' \in 2^A_{k,q}$.
First suppose that $W \in F(p,k,q)$. To prove that $W \in 2^A_{k,q}$, suppose on the contrary that $W \notin 2^A_{k,q}$. Then, for some $j_0 \in [l]$, either $|A_{j_0}| \geq q_{j_0}$ and $|W^{j_0}| < q_{j_0}$, or $|A_{j_0}| < q_{j_0}$ and

[7]Two committees have the same diversity structure if they both contain the same number of candidates from each class.

$A_{j_0} \not\subset W$. In both cases, since $\sum_{j=1}^{l} q_j \leq k$, there exists $j_1 \in [l]$ such that $|W^{j_1}| > q_{j_1}$. Let $a \in W^{j_1}$ and $b \in A_{j_0} \setminus W^{j_0}$ (such a candidate b exists by assumption). We set $T = W \setminus \{a\}$ and $W' = T \cup \{b\}$. It is clear that $\delta^{j_1}(W') = \delta^{j_1}(W) = q_{j_1}$ and $|W^{j_0}| < |W'^{j_0}| \leq q_{j_0}$. Thus, the diverse score of W' is equal to

$$DS_{f_{m,k}}(p, q, W') = S_{f_{m,k}}\Big(p, T \cup \{b\}\Big) + \sum_{j \neq j_0} \Big(\delta^j(W)M\Big) + |W'^{j_0}|M.$$

Therefore, the difference of diverse scores between W' and W is equal to

$$DS_{f_{m,k}}(p, q, W') - DS_{f_{m,k}}(p, q, W) = \Big(|W'^{j_0}| - |W^{j_0}|\Big)M - \Big(S_{f_{m,k}}(p, W) - S_{f_{m,k}}(p, W')\Big).$$

Since $|W'^{j_0}| = |W^{j_0}| + 1$, it follows that

$$DS_{f_{m,k}}(p, q, W') - DS_{f_{m,k}}(p, q, W) = M - \Big(S_{f_{m,k}}(p, W) - S_{f_{m,k}}(p, W')\Big). \qquad (12)$$

By assumption, M satisfies (11). This implies that

$$DS_{f_{m,k}}(p, q, W') - DS_{f_{m,k}}(p, q, W) > \Big(nf_{m,k}(I_0) - S_{f_{m,k}}(p, W)\Big) + \Big(S_{f_{m,k}}(p, W') - nf_{m,k}(J_0)\Big) \geq 0.$$

A contradiction holds since $W \in F(p, k, q)$. Hence, for all $j \in [l]$, either $|A_j| \geq q_j$ and $|W^j| \geq q_j$, or $|A_j| < q_j$ and $A_j \subset W$; that is, $W \in 2^A_{k,q}$.
Now consider any other committee $W' \in 2^A_{k,q}$. Then, $\delta^j(W) = \delta^j(W')$ for all $j \in [l]$ and, thus, the two committees W and W' receive the same total diversity reward. Consequently,

$$DS_{f_{m,k}}(p, q, W) - DS_{f_{m,k}}(p, q, W') = S_{f_{m,k}}(p, W) - S_{f_{m,k}}(p, W') \geq 0,$$

since $W \in F(p, k, q)$. Thus, W is a q-diverse committee with the highest excellence score.

Conversely, assume that W is a q-diverse committee with the highest excellence score. Since W has the same total diversity reward with any other q-diverse committee, it holds that $DS_{f_{m,k}}(p, W) \geq DS_{f_{m,k}}(p, W')$ for all $W' \in 2^A_{k,q}$. For any non q-diverse committee $W' \in 2^A_k$, W receives at least on more diversity reward M than W'. Therefore,

$$\begin{aligned} DS_{f_{m,k}}(p, q, W) - DS_{f_{m,k}}(p, q, W') &\geq S_{f_{m,k}}(p, W) - S_{f_{m,k}}(p, W') + M \\ &> [nf_{m,k}(I_0) - S_{f_{m,k}}(p, W')] + [S_{f_{m,k}}(p, W) - nf_{m,k}(J_0)] \geq 0. \end{aligned}$$

This proves that W maximizes the diverse score among committees in 2^A_k and, thus, $W \in F(p, k, q)$. ■

Before we continue, it is worth mentioning that (11) provides a value of $M(f)$ such that any greater diversity reward guarantees the selection of the q-diverse committee(s) with the highest excellence score. This threshold does not depend on the preferences of voters. However, smaller values of the threshold can be obtained even as a function of the profile in consideration. In this case, we can consider a value $M(f, p)$ such that the quantity in (12) is positive; that is,

$$M(f,p) = \max_{T \in 2^A_{k-1}\,;\, a,b \notin T} \Big(S_{f_{m,k}}\big(p, T \cup \{a\}\big) - S_{f_{m,k}}\big(p, T \cup \{b\}\big) \Big). \tag{13}$$

However, since the total diversity reward should only depend on the composition of any committee and not on the preferences of voters, we simply consider a constant reward which can be declared right at the beginning of the selection process independently of the preferences of voters. Moreover, since our main concern is to select a committee that is closest to the diversity constraint, Theorem 1 suggests to consider from now the threshold $M(f)$ provided in (11).

Definition 9 *A DCSR F is a diverse committee scoring rule if there exists a family of scoring functions $f = (f_{m,k})_{k \leq m-1}$ and a diversity reward $M > M(f) = n\big(f_{m,k}(I_0) - f_{m,k}(J_0)\big)$ such that, for any triplet $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, $F(p, k, q)$ is the set of committee(s) that maximize(s) the diverse score with respect to the couple (M, f).*

It is worth mentioning that the diverse score allows us to obtain a complete ranking of all the committees, and once the diversity reward M is sufficiently large, any best committee in the derived complete ranking is necessarily a q-diverse committee with the highest excellence score as shown in Theorem (1). However, using the diverse score has the practical advantage of applying the voting process in one step, rather than in two steps by first determining all the q-diverse committees and then choosing among them based on their excellence scores.

In the sequel, when the committee size k is fixed for a given number of candidates m, we will simply denote by f the committee scoring function $f_{m,k}$ and by α the scoring vector α^k to ease notations.

4 Desirable axioms for selecting a diverse committee

In this section, we define some properties that we consider to be suitable in the diverse committee selection framework. Some of the properties that we present in this paper are adjustments of well-known properties defined in the literature for the general multiwinner setting, whereas others are proper to our model.

4.1 Optimal diversity and Pareto envy-freeness

We start by providing the properties that are proper to the diverse committee selection setting and which cannot make sense in the general setting. The first properties that we focus on in this category aim to guarantee the main goal previously emphasized, namely the combination of both the excellence in terms of scores and diversity. They are obtained by reshaping two axioms, called *type optimality* and *justified envy-freeness* that have been proposed by Aziz (2019) in order to find a committee that is as close as possible to satisfying the diversity constraint while also selecting the best candidates.

Roughly speaking, a committee is said to be type optimal if it is as close as possible to the diversity constraint. More exactly, a committee W is type optimal if by replacing any candidate in W by any other candidate outside W we cannot obtain a committee that is closer to the constraint than W. Due to the partition of the set of candidates into disjoint classes in our model, we propose the following variant of type optimality, which we call *optimal diversity.*

Definition 10 *A DCSR F satisfies optimal diversity if, for any triplet $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, it holds that $|W^j| \geq q_j$ for all $W \in F(p, k, q)$ and all $j \in [l]$ whenever $|A_j| \geq q_j$ and $A_j \subset W$ otherwise.*

In other words, a DCSR satisfies optimal diversity if it always selects a committee that respects the quota for every class with enough candidates to achieve the quota, and contains all the candidates from any class with a number of candidates less than the enforced quota bound.

Proposition 1 *Every diverse committee scoring rule satisfies optimal diversity.*

Proof. It follows from Theorem 1 since $F(p, k, q) \subseteq 2^A_{k,q}$ and any q-diverse committee satisfies the conditions required by optimal diversity. ■

We can immediately deduce the following corollary.

Corollary 1 *Let F be a diverse committee scoring rule. For any triplet (p, k, q), if there exists at least one size-k committee satisfying the diversity constraint q, then every committee in $F(p, k, q)$ satisfies the diversity constraint.*

Proof. It follows immediately from Proposition 1 since $2^A_{k,q}$ coincides in this case with the set of all committees satisfying the diversity constraint. ■

Note that optimal diversity is desirable to ensure a certain degree of diversity in the selected committee, but it does not take into account the excellence of candidates. By adapting a common property in matching theory,[8] Aziz (2019) defined justified envy-freeness for this purpose, which addresses the following idea: If a candidate a is ranked higher than a candidate b in the arbitrary (social) weak order over the set of candidates, the only reason one could leave a and choose b is precisely the proximity to the constraint q. We reshape this axiom to make it suitable for our framework and we call it *Pareto envy-freeness.*

Definition 11 *A DCSR F satisfies Pareto envy-freeness if for all triplets $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, and for any two candidates a and b such that $b \succ_i a$ for all $i \in N$, it holds that if $a \in W$ for some $W \in F(p, k, q)$ and $|W^{j(a)}| > q_{j(a)}$, then $b \in W'$ for some $W' \in F(p, k, q)$.*

[8]See Ehlers et al. (2014), Goto et al. (2017), and Kamada and Kojima (2015), among others.

Intuitively, Pareto envy-freeness requires that when a candidate a belongs to a winning committee wherein the class of a is already overrepresented, it holds that if all individuals (and thus the society) prefer a candidate b to candidate a then b is also selected in at least one winning committee. Note that in accordance with the spirit of justified envy-freeness, our goal is to consider a social outcome where candidate b is ranked higher than candidate a. In order to ensure that b will be collectively better ranked than a for any reasonable rule that we use to aggregate individual preferences, the only minimal requirement is to consider that all the voters prefer b to a.

Proposition 2 *Every diverse committee scoring rule satisfies Pareto envy-freeness.*

Proof. Let F be a diverse committee scoring rule based on a committee scoring function f. Let p be a preference profile, $k \in [m-1]$ be the committee size, and q be a diversity constraint. Let $a, b \in A$ such that $b \succ_i a$ for every $i \in N$. Assume that $a \in W$ for some $W \in F(p,k,q)$ with $|W^{j(a)}| > q_{j(a)}$. By optimal diversity, either $|W^{j(b)}| \geq q_{j(b)}$ or $A_{j(b)} \subset W$ hold. In the latter case, it holds that $b \in W$ and $W' = W \in F(p,k,q)$. In the former case, it can also be the case that $b \in W \in F(p,k,q)$; otherwise, let $W' = \Big(W \setminus \{a\}\Big) \cup \{b\}$. It holds that $\delta^{j(b)}(W') = \delta^{j(b)}(W) = q_{j(b)}$ and $\delta^{j(a)}(W') = \delta^{j(a)}(W) = q_{j(a)}$. Thus, W and W' both receive the same total diversity reward and the difference of diverse scores between W' and W is simply

$$\begin{aligned} DS_f(p,q,W') - DS_f(p,q,W) &= S_f(p,W') - S_f(p,W) \\ &= \sum_{i \in N} f\Big(r(p_i,W')\Big) - \sum_{i \in N} f\Big(r(p_i,W)\Big) \\ &\geq 0, \end{aligned}$$

since $r(p_i,W') \succeq r(p_i,W)$ for all $i \in N$ and then $f\Big(r(p_i,W')\Big) \geq f\Big(r(p_i,W)\Big)$ for all $i \in N$. Hence, $W' \in F(p,k,q)$ and $b \in W'$. ■

4.2 Constraint fairness and type independence

Let us now add some further notations. For any diversity constraint q and for any $j \in [l]$, we denote by (q_{-j}, q'_j) the diversity constraint obtained from q by replacing q_j by q'_j. For any profile p, we simply write $F(p,k)$ as the outcome of the committee selection rule F with the null diversity constraint $\big($i.e., $q_j = 0$ for all $j \in [l]\big)$, meaning that there is no diversity constraint in the committee selection process.

Another desirable property which is also proper to our setting is called *constraint fairness.*

Definition 12 *A DCSR F satisfies constraint fairness if, for any two candidates $a, b \in A$ such that $j(a) = j(b)$, if $a \in W$ for some $W \in F(p,k)$ and $b \notin W$ for all $W \in F(p,k)$, then for any diversity constraint q and for any $W' \in F(p,k,q)$, $b \in W'$ implies $a \in W'$.*

Roughly speaking, a DCSR F satisfies the constraint fairness property if introducing the diversity constraint does not foster a candidate having a given type while penalizing another candidate from the same type.

Proposition 3 *Every diverse (weakly) separable committee scoring rule satisfies constraint fairness.*

Proof. Let F be a diverse (weakly) separable committee scoring rule defined via a scoring vector α, p be a preference profile, and $k \in [m-1]$ be the committee size. Let $a, b \in A$ be two candidates such that $j(a) = j(b)$, $a \in W$ for some $W \in F(p,k)$, and $b \notin W$ for all $W \in F(p,k)$. Then, it holds that $S_\alpha(p,a) > S_\alpha(p,b)$ by the definition of $F(p,k)$. Let q be a diversity constraint. Assume by contradiction that there is a committee $W' \in F(p,k,q)$ such that $b \in W'$ and $a \notin W'$. Consider the committee W'' obtained from W' by replacing b by a. Then, $W'' = \Big(W' \setminus \{b\}\Big) \cup \{a\}$ and it follows that W' and W'' have the same diversity structure since $j(a) = j(b)$. Thus, $|W''^j| = |W'^j| \geq \delta(A_j)$ (by Theorem 1) for all $j \in [l]$ and $W', W'' \in 2^A_{k,q}$. But W'' and W' satisfy

$$DS_\alpha(p,q,W'') - DS_\alpha(p,q,W') = \sum_{x \in W''} S_\alpha(p,x) - \sum_{x \in W'} S_\alpha(p,x) = S_\alpha(p,a) - S_\alpha(p,b) > 0.$$

This is a contradiction since $W' \in F(p,k,q)$ and one should have $DS_\alpha(p,q,W'') - DS_\alpha(p,q,W') \leq 0$. Hence, for any diversity constraint q and for any $W' \in F(p,k,q)$, $b \in W'$ implies $a \in W'$. ■

The following example shows that constraint fairness is not satisfied by all diverse committee scoring rules.

Example 2 *Consider a set of candidates $A = \{a_1, a_2, a_3, a_4, b_1, b_2\}$ partitioned according to gender (for instance) such that $A_1 = \{a_{,1}, a_2, a_3, a_4\}$ is the class of men and $A_2 = \{b_1, b_2\}$ is the class of women. Assume that $k = 2$ and $q = (0, 1)$. Consider the following preference profile*

$$p = \begin{bmatrix} a_1 & a_1 & a_2 & a_2 \\ a_3 & a_3 & a_3 & a_3 \\ a_4 & a_2 & a_4 & a_1 \\ a_2 & a_4 & a_1 & a_4 \\ b_1 & b_1 & b_1 & b_1 \\ b_2 & b_2 & b_2 & b_2 \end{bmatrix}$$

Under the version of the Chamberlin-Courant rule that has been presented in (9), the only winning committee without any diversity constraint is $\{a_1, a_2\}$ with a total score of 20. Now, it can be checked that enforcing the diversity constraint results in the winning committees $\{a_3, b_1\}$ and $\{a_3, b_2\}$ under the diverse Chamberlin-Courant rule with a maximal diverse score of $16 + M$, where the diversity reward satisfies $M > 16$. Thus, the property of constraint fairness is not satisfied since a_1 is selected whereas a_3 is not if we do not con-

sider the diversity constraint, but enforcing the diversity constraint results in the selection of a_3 and not a_1.

The next property that we consider as suitable for the diverse committee selection setting requires the attribute classes to be independent. Intuitively, the attribute classes are independent if the quality of the members from a given class in any preference profile does not affect (upward or downward) that of candidates from any other class. In other words, if some candidates of a given type end up in a winning committee it is because those candidates deserve it, not because they are taking advantage of the potential of candidates from other types. We call this property *type independence*.

Definition 13 *A DCSR F satisfies type independence if for every triplet $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$, for every $j \in [l]$ and for any two committees W and W' such that $W^j = W'^j$, if $W \in F(p, k, q)$ and $W' \notin F(p, k, q)$, then for every subset $B_j \subseteq A_j$ with $|B_j| = |W'^j|$ it holds that $W'' = (W' \setminus W'^j) \cup B_j \notin F(p, k, q)$.*

Proposition 4 *Every diverse (weakly) separable committee scoring rule satisfies type independence.*

Proof. Let p be a preference profile, k be the committee size, and q be a diversity constraint. Let F be a diverse (weakly) separable committee scoring rule associated with the scoring vector α. Let $j \in [l]$ be a type and W and W' be two size-k committees such that $W \in F(p, k, q)$, $W' \notin F(p, k, q)$, and $W^j = W'^j$. Consider $B_j \subseteq A_j$ with $|B_j| = |W^j| = |W'^j|$ and let $W'' = (W' \setminus W'^j) \cup B_j$. Note that W' and W'' have the same diversity structure. There are two possible cases. First, suppose that $W' \notin 2^A_{k,q}$. It follows that $W'' \notin 2^A_{k,q}$. Therefore, $W'' \notin F(p, k, q)$ by Theorem 1. Now, suppose that $W' \in 2^A_{k,q}$. In this case, W and W' have the same total diversity reward and, thus, $S_\alpha(p, W') < S_\alpha(p, W)$. This implies that

$$
\begin{aligned}
S_\alpha(p, W') = \sum_{a \in W'} S_\alpha(p, a) &= \sum_{a \in W'^j} S_\alpha(p, a) + \sum_{a \in W' \setminus W'^j} S_\alpha(p, a) \\
&< S_\alpha(p, W) \\
&= \sum_{a \in W^j} S_\alpha(p, a) + \sum_{a \in W \setminus W^j} S_\alpha(p, a).
\end{aligned}
$$

Since $W^j = W'^j$, the following holds

$$
\sum_{a \in W' \setminus W'^j} S_\alpha(p, a) \quad < \quad \sum_{a \in W \setminus W^j} S_\alpha(p, a).
$$

Then, the committee $W'' = (W' \setminus W'^j) \cup B_j$ satisfies

$$\begin{aligned} S_\alpha\Big(p, W''\Big) &= \sum_{a \in B_j} S_\alpha(p,a) + \sum_{a \in W' \setminus W'^j} S_\alpha(p,a) \\ &< \sum_{a \in B_j} S_\alpha(p,a) + \sum_{a \in W \setminus W^j} S_\alpha(p,a) \\ &= S_\alpha\Big(p, (W \setminus W^j) \cup B_j\Big) \\ &\leq S_\alpha(p, W). \end{aligned}$$

Hence, $S_\alpha\Big(p, W''\Big) < S_\alpha(p, W)$ and $W'' \notin F(p,k,q)$. ■

The following example shows that type independence is not satisfied by every diverse committee scoring rule. We take the diverse Chamberlin-Courant rule as an example.

Example 3 *Consider a set of candidates $A = \{a_1, a_2, a_3, b_1, b_2\}$ partitioned according to gender (for instance) such that $A_1 = \{a_1, a_2, a_3\}$ is the class of men and $A_2 = \{b_1, b_2\}$ is the class of women. Assume that $k = 2$ and $q = (1,1)$. Consider the following preference profile with four voters*

$$p = \begin{bmatrix} b_1 & b_1 & b_2 & b_2 \\ a_1 & a_2 & a_2 & a_3 \\ a_2 & a_1 & a_1 & a_2 \\ b_2 & b_2 & b_1 & b_1 \\ a_3 & a_3 & a_3 & a_1 \end{bmatrix}$$

Under the diverse Chamberlin-Courant rule, it can be checked that the committee $\{a_1, b_2\}$ is a winning committee with the maximal diverse score of $13 + 2M$ (with $M > 12$) whereas the committee $\{a_1, b_1\}$ is a losing committee with the diverse score of $11 + 2M$. Moreover, if we replace a_1 by a_2 into the latter committee, we obtain $\{a_2, b_1\}$ which is also a winning committee with the same maximal diverse score of $13 + 2M$. This proves that the diverse Chamberlin-Courant rule fails to satisfy type independence.

4.3 Further properties

We now provide some properties that are adjustments of some well-known properties of multiwinner voting rules that have been defined for the general setting. Clearly, not all axioms known in the literature of committee selection can be adapted to the diversity framework. We will therefore focus on those axioms that make sense in our framework and that seem to us as important to be satisfied. This point will be discussed more thoroughly at the end of this section.

The *weak unanimity* property has been defined by Elkind et al. (2017) as follows: For every preference profile p and every integer $k \in [m-1]$, if there is a committee $W \in 2^A_k$ such that each voter ranks the k candidates from W on top, then $W \in F(p,k)$. From the above definition, we give here a variant of weak unanimity for diverse committee elections that we call *weak restricted unanimity.*

Definition 14 *A DCSR F satisfies weak restricted unanimity if for all triplets $(p, k, q) \in \mathcal{P}^n \times [m-1] \times \mathbb{N}^l$ and every $j \in [l]^*_q$, if there is a subset $X_j \subseteq A_j$ of size $x_j \leq q_j$ such that each voter ranks all the x_j candidates from X_j before every other candidate from $A_j \setminus X_j$, it holds that $X_j \subseteq W$ for some $W \in F(p, k, q)$.*

Intuitively, the condition of weak restricted unanimity requires that when a group of candidates of the same type, and of cardinality less than or equal to the required diversity quota for that type, is ranked by all the voters before every other candidate of that type, then this group of candidates must necessarily belong to a winning committee.

Proposition 5 *Every diverse committee scoring rule satisfies weak restricted unanimity.*

Proof. Let F be a diverse committee scoring rule associated to a family of scoring functions f and a diversity reward M, p be a preference profile, $k \in [m-1]$ be the committee size, and q be a diversity constraint. Let $j \in [l]^*_q$ and $X_j \subseteq A_j$ be a subset of A_j of size $x_j \leq q_j$ such that each voter ranks all the candidates from X_j before every candidate from $A_j \setminus X_j$. Assume by contradiction that there is no winning committee W such that $X_j \subseteq W$, which means that the diverse score DS_f of any committee containing X_j is (strictly) less than the diverse score DS_f of any winning committee. Let $W \in F(p, k, q)$. By optimal diversity, it follows that $|W^j| \geq x_j$ and then there are some candidates $a_1, \ldots, a_s \in A_j \setminus X_j$ with $1 \leq s \leq x_j$ such that $a_t \in W$ for all $t \leq s$, and there are some candidates $b_1, \ldots, b_s \in X_j$ such that $b_t \notin W$ for all $t \leq s$. Let us set $W' = \big(W \setminus \{a_1, \ldots, a_s\}\big) \cup \{b_1, \ldots, b_s\}$. Since all the candidates b_t are ranked higher than all the candidates a_t, then it follows that $r(p_i, W') \succeq r(p_i, W)$ for all $i \in N$ and $f\Big(r(p_i, W')\Big) \geq f\Big(r(p_i, W)\Big)$ for all $i \in N$. This implies that $S_f(p, W') \geq S_f(p, W)$. Moreover, since $\delta^j(W') = \delta^j(W)$ for all $j \in [l]$, W and W' receive the same total diversity reward. The difference of diverse scores between W' and W is then equal to

$$DS_f(p, q, W') - DS_f(p, q, W) = S_f(p, W') - S_f(p, W) \geq 0.$$

This inequality means that W' has a diverse score greater than or equal to that of W, which is a contradiction with the assumption. Hence X_j is contained in at least one winning committee. ∎

Kamwa (2017) defined the *Pareto criterion* in order to study the behavior of stable multiwinner voting rules.[9] In the context of committee elections, the Pareto criterion requires that if there are two candidates a and b such that a is always ranked ahead of b, then if there is a winning committee that includes b, there is also one that includes a. Formally, the Pareto criterion is satisfied if, for every preference profile p, every committee size $k \in [m-1]$, and any two candidates a and b such that $a \succ_i b$ for all $i \in N$, it holds

[9] A committee selection rule is stable if it always picks a *weak Condorcet committee* when such a committee exists. A weak Condorcet committee is a committee such that no member of that committee is defeated in pairwise majority comparisons by a candidate that does not belong to it (see, for instance, Bubboloni et al., 2020; Diss and Doghmi, 2016; Diss et al., 2020; Diss and Mahajne, 2020; Gehrlein, 1985).

that, $b \in W$ for some $W \in F(p,k)$ implies $a \in W'$ for some $W' \in F(p,k)$. We reshape this property in the diverse committee selection framework by naturally making restrictions on different classes of the set of candidates A. We call this property the *restricted Pareto criterion*.

Definition 15 *A DCSR F satisfies the restricted Pareto criterion if for each preference profile p, each $k \in [m-1]$, any $j \in [l]$, and any two candidates a and b such that $j(a) = j(b)$, it holds that if $a \succ_i b$ for all $i \in N$, then for any diversity constraint q, $b \in W$ for some $W \in F(p,k,q)$ implies that $a \in W'$ for some $W' \in F(p,k,q)$.*

Intuitively, this axiom requires that when all voters unanimously rank a candidate a above another candidate b of the same type, it holds that if candidate b belongs to a given winning committee, then this must also be the case for candidate a.

Proposition 6 *Every diverse committee scoring rule satisfies restricted Pareto criterion.*

Proof. Let p be a preference profile. Consider any two candidates a and b such that $j(a) = j(b)$. Assume that $a \succ_i b$ for all $i \in N$. Let $W \in F(p,k,q)$ be a winning committee such that $b \in W$. Either $a \in W$ or $a \notin W$. In the latter case, the committee $W' = (W \setminus \{b\}) \cup \{a\}$ has the same diversity structure as W and, then, the two committees receive the same total diversity reward. Moreover, we have $r(p_i, W') \succeq r(p_i, W)$ for all $i \in N$ since a is ranked higher than b by all the voters. This implies that $S_f(p, W') \geq S_f(p, W)$. Therefore, $DS_f(p,q,W') - DS_f(p,q,W) = S_f(p,W') - S_f(p,W) \geq 0$. Thus, W' is a winning committee that contains a. ∎

Following Elkind et al. (2017), a committee selection rule F satisfies *committee monotonicity* if, for every preference profile p, the following conditions hold: (1) For each $k \in [m-1]$, if $W \in F(p,k)$, then there exists a committee $W' \in F(p,k+1)$ such that $W \subseteq W'$; (2) for each $k \in [m-1]$, if $W \in F(p,k+1)$, then there exists $W' \in F(p,k)$ such that $W' \subseteq W$. Note that the committee monotonicity property has been widely discussed by many other authors in the context of multiwinner elections under different names (see, for instance, Barberà and Coelho, 2008; Kilgour and Marshall, 2012; Ratliff, 2003; Staring, 1986). A variant of this property has also been studied by Kamwa and Merlin (2015). In our setup, we propose a property that is similar to committee monotonicity which we call *constraint monotonicity*. Our proposal does not depend on the committee size but on the diversity constraint.

Definition 16 *A DCSR F satisfies constraint monotonicity if for every preference profile p, each $k \in [m-1]$, and any diversity constraint q such that $\sum_{j=1}^{l} q_j < k$, it holds that: (1) For every $j \in [l]$, if $W \in F(p,k,q)$, then there exists $W' \in F\big(p,k,(q_{-j}, q_j+1)\big)$ such that $W^j \subseteq W'^j$; (2) for every $j \in [l]$, if $W' \in F\big(p,k,(q_{-j}, q_j+1)\big)$, then there exists $W \in F(p,k,q)$ such that $W^j \subseteq W'^j$.*

Basically, constraint monotonicity stipulates that when the required diversity quota increases for a given type none of the already selected candidates of that type should be dropped. In addition, lowering the diversity quota for a given type must not lead to the selection of completely new candidates of the same type. Note that what is modified in the definition of committee monotonicity is the committee size k, but in that of constraint monotonicity, we are concerned with the diversity quota q_j of type $j \in [l]$ while the committee size remains fixed.

Note that Elkind et al. (2017) used committee monotonicity to characterize a specific class of committee selection rules called the *best-k rules*. Before giving the definition of best-k rules, let us remind the reader that a social preference function g is a function assigning to each preference profile p, the set $g(p)$ of (tied) linear orders over the set of candidates A. According to a best-k rule, we sort the candidates according to a linear order provided by a social preference function and pick the top k ones. The formal definition is given as follows:

Definition 17 *A committee selection rule is a best-k rule if there exists a social preference function g such that for any preference profile p and each $k \in [m-1]$, it holds that $W \in F(p,k)$ if and only if there is a linear order $\succ \in g(p)$ that ranks the candidates in W in the top positions; that is, $a \succ b$ for all $a \in W$ and $b \in A \setminus W$.*

We can see that the class of best-k rules is incompatible with the spirit of electing a diverse committee since selecting the top-k candidates of the ranking generated by the social preference function might fail to guarantee the diversity requirement. We provide below a suitable class of diverse committee selection rules equivalent to the class of best-k rules, which depend on the attribute classes. We call them the *best-top diverse* rules.

Definition 18 *A DCSR F is a best-top diverse rule if for each $k \in [m-1]$, there exists a social preference function g_k such that for any preference profile p and any diversity constraint q, it holds that $W \in F(p,k,q)$ if and only if there is a linear order $\succ \in g_k(p)$ such that for every $j \in [l]$, $\succ^j$ ranks the candidates in W^j on the top positions; that is, $a \succ^j b$ for all $a \in W^j$ and $b \in A_j \setminus W^j$.*

In other words, we here sort all the candidates of the same type according to their positions in the social ranking and pick the top ones. Recall that $\succ^j$ is the restricted linear ranking induced by the linear order $\succ$ on the class A_j.

Theorem 2 *A DCSR is a best-top diverse rule if and only if it is constraint monotonic.*

Proof. Let F be a best-top diverse committee selection rule, p be a preference profile, $k \in [m-1]$ be the committee size, and q be a diversity constraint such that $\sum_{j=1}^{l} q_j < k$.[10]
(1) Let $W \in F(p,k,q)$. Then, there exists a linear order $\succ \in g_k(p)$ such that for each $j \in [l]$, W^j consists of the top candidates of $\succ^j$. Let $j \in [l]$.

[10] The sum of the diversity quotas of all types should be strictly smaller than k in order to be able to increase a quota of type j by 1 following the definition of constraint monotonicity.

- If $|W^j| \geq q_j + 1$, then W is also a winning committee for the diversity constraint $q' = (q_{-j}, q_j + 1)$ and the result holds.
- If $|W^j| \leq q_j$, since $\sum_{t=1}^{l} q_t < k$, then there exists $j_0 \in [l]$ such that $|W^{j_0}| > q_{j_0}$. Let $a \in W^{j_0}$ such that $x \succ^{j_0} a$ for all $x \in W^{j_0} \setminus \{a\}$, and $b \in A_j \setminus W^j$ such that $b \succ^j y$ for all $y \in A_j \setminus W^j$. In other words, a is the worst-ranked candidate from W^{j_0} with respect to $\succ^{j_0}$ and b is the best-ranked candidate from $A_j \setminus W^j$ with respect to $\succ^j$. By letting $W' = \Big(W \setminus \{a\}\Big) \cup \{b\}$, it follows that $W' \in F\big(p, k, (q_{-j}, q_j + 1)\big)$ and $W^j \subset W'^j$.

(2) Let $j \in [l]$ and $W' \in F(p, k, q')$ with $q' = (q_{-j}, q_j + 1)$. Then, there is a linear order $\succ \in g_k(p)$ on A such that the candidates from W'^j are in the top positions of $\succ^j$. Two cases are possible:

- **Case 1:** If $|W'^j| > q_j + 1$, then W' is still winning for the diversity constraint q and the result holds.
- **Case 2:** If $|W'^j| \leq q_j + 1$, W' can still be a winning committee for the diversity constraint q and the result holds directly. Otherwise, let $a \in W'^j$ such that $x \succ^j a$ for all $x \in W'^j \backslash \{a\}$. Then there is a committee $W \in F(p, k, q)$ such that $W^j = W'^j \backslash \{a\}$, and then $W^j \subset W'^j$.

Conversely, assume that F satisfies constraint monotonicity. We have to show that F is a best-top diverse rule. Let $k \in [m-1]$ be the given committee size and p be a preference profile. Assume that the set of candidates is $A = \{a_1, \ldots, a_m\}$, and for any class A_j we can write $A_j = \{a_1^j, \ldots, a_{m_j}^j\}$. Let $g_k(p)$ be the set of all linear orders $\succ$ on A such that for every $j \in [l]$, if the restriction $\succ^j$ is $a_{\pi(1)}^j \succ^j \cdots \succ^j a_{\pi(m_j)}^j$ (where π is the permutation which gives the order of the candidates of A_j in the linear order $\succ$), then for any q_{-j} and for all $s \geq 1$, $W_s^j = \{a_{\pi(1)}^j, \ldots, a_{\pi(s)}^j\} \subset W^j$ for some $W \in F(p, k, (q_{-j}, s))$. Such a linear order always exists from the two conditions of constraint monotonicity, and it holds that F is a best-top diverse rule with the social preference function that assigns to each preference profile p the set of linear orders $g_k(p)$ for each $k \in [m-1]$. ■

The following proposition states that the set of best-top diverse rules includes the subclass of diverse (weakly) separable committee scoring rules.

Proposition 7 *Every diverse (weakly) separable committee scoring rule is a best-top diverse rule.*

Proof. Let F be a diverse (weakly) separable committee scoring rule with respect to the family of scoring vectors $\alpha = (\alpha^k)_{k \leq m-1}$. Given a committee size k, let g_k be the social preference function that associates each preference profile p with the set $g_k(p)$ of linear orders on A such that for all $\succ \in g_k(p)$ and any two candidates a and b, $a \succ b$ implies $S_\alpha(p, a) \geq S_\alpha(p, b)$. Let $j \in [l]$ and $a, b \in A_j$ such that $a \succ^j b$. Assume that $b \in W$ for some

$W \in F(p, k, q)$. Then, either $a \in W$ or $a \notin W$. In the latter case, let $W' = \Big(W \setminus \{b\}\Big) \cup \{a\}$. It holds that W and W' earn the same total diversity reward since a and b belong to the same class, and then $DS_\alpha(p, W') - DS_\alpha(p, W) = S_\alpha(p, a) - S_\alpha(p, b) \geq 0$. So, W' is also a winning committee that contains a. Thus, F is a best-top diverse rule associated with the social preference function g_k for each $k \in [m-1]$. ■

We deduce the following corollary from Theorem 2 and Proposition 7.

Corollary 2 *Every diverse (weakly) separable committee scoring rule is constraint monotonic.*

We provide below a simple example to show that not all diverse committee scoring rules are best-top rules.

Example 4 *Consider a set of candidates $A = \{a, b, c, d\}$ partitioned into two classes A_1 and A_2 such that $A_1 = \{a, b, c\}$ and $A_2 = \{d\}$. Assume that $k = 2$ and $q = (1, 0)$. Consider the following preference profile with five voters*

$$p = \begin{bmatrix} a & d & d & d & d \\ b & b & c & c & b \\ c & c & b & b & c \\ d & a & a & a & a \end{bmatrix}$$

Under the diverse Chamberlin-Courant rule, it can be checked that the committee $\{a, d\}$ is the only winning committee with the maximal diverse score of $15 + M$ (with $M > 10$). However, it can be checked that the only winning committee for the diversity constraint $q' = (2, 0)$ is the committee $\{b, c\}$ with the maximal diverse score of $10 + 2M$. Thus, candidate a is selected into the winning committee for the quota $q_1 = 1$ while increasing the quota q_1 to 2 leads to the selection of b and c and not a. This proves that the diverse Chamberlin-Courant rule fails to satisfy constraint monotonicity and then it is not a best-top diverse rule.

As argued by Elkind et al. (2017), separable committee scoring rules belong to the class of best-k rules, which is not the case for weakly separable committee scoring rules. This is due to the fact that the social ranking outputted by a weakly separable scoring rule depends on the committee size and, then, the transition from k to $k+1$ might change the social ranking. In other words, the conditions of committee monotonicity would be failed. However, this issue is released for the best-top diverse rules since in the definition of constraint monotonicity, the quota of a given class changes while the committee size does not. This allows diverse weakly separable committee scoring rules to also be best-top diverse rules.

Let us now move to our next property. Note that Elkind et al. (2017) and Faliszewski et al. (2019) characterized the (weakly) separable committee scoring rules as the only committee scoring rules that satisfy the non-crossing monotonicity property. Roughly speaking, a committee selection rule satisfies the non-crossing monotonicity property if, by shifting a candidate a from a winning committee one position forward in one individual preference, that committee stays winning if the prior candidate ranked immediately before a in that preference does not belong to that committee. For our framework, we provide below a variant of the non-crossing monotonicity property that we call *type monotonicity.*

Definition 19 *A DCSR F satisfies type monotonicity if for every profile p, every $k \in [m-1]$, every diversity constraint q, and any candidate $a \in A$ such that $a \in W$ for some $W \in F(p,k,q)$, then for any preference profile p' obtained from p by shifting the candidate a one position forward in some individual preference p_i, it holds that if candidate a was ranked immediately below some candidate b such that $j(a) \neq j(b)$, then $W^{j(a)} \subseteq W'^{j(a)}$ for some $W' \in F(p',k,q)$.*

Intuitively, if a is shifted one position forward after a swap with a candidate b from a different class, then all the candidates from the same class as a that were selected in W should still be selected in a winning committee when we consider the new profile. In contrast with the original form of non-crossing monotonicity, we do not require the prior candidate b to be not selected.

Theorem 3 *A diverse committee scoring rule F is (weakly) separable if and only if it satisfies type monotonicity.*

Proof. Let F be a diverse (weakly) separable committee scoring rule defined by the family of scoring vector $\alpha = (\alpha^k)_{k \leq m-1}$. Let a be a candidate belonging to a winning committee $W \in F(p,k,q)$ and i be a voter that ranks candidate a at a given position $t \in \{2, \cdots, m\}$. Let b be the candidate ranked at the position $(t-1)$ in $\succ_i$ such that $j(b) \neq j(a)$. After voter i swaps a and b, candidate a gains $\alpha_{t-1} - \alpha_t$ points, candidate b loses $\alpha_{t-1} - \alpha_t$ points, and all the other candidates keep the same individual score. Since $b \notin A_{j(a)}$, it holds that the swapping does not decrease the total score of the candidates from $W^{j(a)}$. Therefore, all the members from $W^{j(a)}$ still belong to a winning committee with respect to the new profile p'.
Conversely, let F be a diverse committee scoring rule associated with a family of scoring functions f and a reward M. Assume that F satisfies type monotonicity and let us show that F is weakly separable.
Consider the preference profile p with $m!$ voters wherein each linear order is cast by one voter. In this case, all size-k committees have the same excellence score and the winning committee is any size-k committee with the maximal diverse score. Let $I = (i_1, \ldots, i_k)$ be the rank of such a committee that contains a position $t \in \{2, \ldots, m\}$. Let $i \in N$ be a given voter such that the two candidates that fill the positions $t-1$ (candidate b) and t (candidate a) are from different classes, and let $W(I)$ be the set of k candidates such that $r(p_i, W(I)) = I$. It holds that $W(I)$ is a winning committee with respect to p. Let p' be the profile obtained after swapping candidate b at the position t and candidate

a at the position $t-1$ in the linear order of voter i and let I' be the committee rank obtained from I after this permutation. Let us note that $I = I'$ if I contains the position $t-1$; that is, if candidate b also belongs to $W(I)$. Since F is type monotonic, we have $W(I)^{j(a)} \subseteq W$ for some $W \in F(p', k, q)$. Let $J = r(p_i, W)$ be the rank of the committee W in the linear order of voter i, and let $J' = r(p'_i, W)$; obviously, J contains position t and J' contains position $t-1$. The difference of scores of W between the two profiles p' and p is equal to $f_{m,k}(J') - f_{m,k}(J)$. This difference of score does not depend on the rank I. Indeed, let I^* be another rank of a committee maximizing the diversity score, J^* and $J^{*\prime}$ be the ranks obtained from I^* analogously, and W^* be the corresponding committee (the committee such that $r(p_i, W^*) = J^*$). It holds that W and W^* have the same maximal score across the two profiles p and p'; it then follows that $f_{m,k}(J') - f_{m,k}(J) = f_{m,k}(J^{*\prime}) - f_{m,k}(J^*)$. Thus, we can define the function $h_{m,k} : \{2, \ldots, m\} \to \mathbb{R}_+$ such that $h_{m,k}(t-1) = f_{m,k}(J') - f_{m,k}(J)$ for any rank J containing the position t.
The rest of the proof is similar to that of Theorem 10 in Faliszewski et al. (2019) and we point the reader toward that work. ■

Another desirable property of multiwinner elections defined by Elkind et al. (2017) is the *consensus committee* property. This property is originally defined as follows: Let p be a preference profile and k be the committee size. If there is a committee W of size k such that each voter ranks some candidate from W first and each member of W is ranked first by $\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$ voters, then $F(p, k) = \{W\}$. We fit this property to our model by focusing on a single class and we call it *type consensus*.

Definition 20 *A DCSR F satisfies type consensus if for every preference profile p, any diversity constraint q, and any type $j \in [l]^*_q$, it holds that if there is a q_j-element subset Q_j of A_j such that each voter ranks some member of Q_j first, and each member of Q_j is ranked first by $\lfloor \frac{n}{q_j} \rfloor$ or $\lceil \frac{n}{q_j} \rceil$ voters, then $Q_j \subseteq W$ for all $W \in F(p, k, q)$.*

We can check that the diverse form of the Chamberlin-Courant rule given in (9) satisfies the type consensus property since any subset Q_j that meets the conditions of Definition (9) necessary has the maximal excellence score whenever such a subset exists. However, not all diverse committee scoring rules satisfy this property. The next proposition gives an overview regarding diverse separable committee scoring rules.

Proposition 8 *Let F be a diverse separable committee scoring rule based on the scoring vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$.*

1. *If F = diverse SNTV, then F satisfies type consensus.*
2. *If $F \neq$ diverse SNTV, then F satisfies type consensus if $\max_{j \in [l]} q_j < \frac{\alpha_1}{\alpha_2}$ whenever the number of voters is sufficiently large.*
3. *If $F \neq$ diverse SNTV, then F fails to satisfy type consensus if there is a type $j \in [l]$ such that $q_j \geq \frac{\alpha_1}{\alpha_2}$.*

Proof. Let F be a diverse separable committee scoring rule based on the scoring vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ and p be a preference profile. Let $q = (q_1, \ldots, q_l)$ be a diversity constraint. Let us prove the three points separately.

1. Assume that F is the diverse SNTV rule; that is $\alpha_2 = 0$. Then each member of Q_j gains at least one point and the score of any non-member of Q_j is null. Since the diversity constraint requires at least q_j members from A_j, then $Q_j \subseteq W$ (by optimal diversity) for all $W \in F(p, k, q)$.

2. Assume that $F \neq$ diverse SNTV and $\max_{j\in[l]} q_j < \frac{\alpha_1}{\alpha_2}$. Let $j \in [l]^*_q$ and Q_j be a q_j-element subset of A_j such that each voter ranks some candidate from Q_j first and each candidate from Q_j is ranked first by $\lfloor \frac{n}{q_j} \rfloor$ or $\lceil \frac{n}{q_j} \rceil$. Then the individual score of every candidate from Q_j is at least $\lfloor \frac{n}{q_j} \rfloor \alpha_1 \geq (\frac{n}{q_j} - 1)\alpha_1 > n\alpha_2$ whenever $n > \frac{q_j \alpha_1}{\alpha_1 - q_j \alpha_2} > 0$ (since $\alpha_1 - q_j\alpha_2 > 0$). Thus, the individual score of each member of Q_j is greater than $n\alpha_2 \geq S(p, x)$ for all $x \in A \setminus Q_j$. Since the selected committee contains at least q_j candidates from A_j, it holds that $Q_j \subseteq W$ for all $W \in F(p, k, q)$. all

3. Assume that $\alpha_2 \neq 0$ and that there is a type $j \in [l]$ such that $q_j \geq \frac{\alpha_1}{\alpha_2}$. Assume that F satisfies the type consensus property. We can construct a profile with m candidates and $q_j + 1$ voters as follows: the class A_j contains at least $q_j + 1$ candidates that we can write $A_j = \{a_1, \ldots, a_{q_j}, a_{q_j+1}, \ldots, a_{m_j}\}$, each of the candidates $a_1, \ldots, a_{q_j - 1}$ is ranked first by one voter, candidate a_{q_j} is ranked first by 2 voters and all the voters rank a_{q_j+1} second. Moreover, from position 3, all the voters rank all the candidates from other classes before all the remaining candidates in A_j and, finally, all the voters who do not rank candidate a_1 first rank it at the last position. By considering a diversity constraint q such that $\sum_{s=1}^{l} q_s = k$, it follows that the selected committee should contain exactly q_j candidates from A_j. Since F satisfies the type consensus property, it follows that $W^j = \{a_1, \ldots, a_{q_j}\}$ for all $W \in F(p, k, q)$, which means that a_1 has a greater individual score than a_{q_j+1}. However, the score of candidate a_{q_j+1} is $n\alpha_2 = (q_j + 1)\alpha_2$ and it follows that $\alpha_1 > (q_j + 1)\alpha_2$ which is a contradiction. Hence F fails the type consensus property. ■

As mentioned earlier in this section, we focused on the properties of multiwinner rules that can be adapted to the diverse committee selection framework. Let us mention that some other well-known properties of multiwinner rules such as *anonymity*, *homogeneity*, *consistency*, and *candidate monotonicity*[11] are trivially satisfied by every diverse committee scoring rule. The anonymity property requires that the result of any election does not depend on the names of voters. Formally, for any triplet (p, k, q) and any permutation π on the set of voters, $F(\pi(p), k, q) = F(p, k, q)$, where $\pi(p) = (p_{\pi(1)}, \ldots, p_{\pi(n)})$. The homogeneity property requires that every duplication of the preferences of voters leads to the same result. Formally, $F(tp, k, q) = F(p, k, q)$ for all $t \in \mathbb{N}^*$, where tp is a new

[11]See, for instance, Elkind et al. (2017).

profile in which each voter's preference in p is replicated the same number of times t. The consistency property requires that if W is a winning committee with respect to two preference profiles p and p' that are respectively cast by two disjoint set of voters N_1 and N_2, then W should still be winning when we merge the two profiles. Formally, if $F(p, k, q) \cap F(p', k, q) \neq \emptyset$, then $F(p + p', k, q) = F(p, k, q) \cap F(p', k, q)$, where $p + p'$ is the profile cast by the set $N_1 \cup N_2$. The property of candidate monotonicity requires that if a candidate a belongs to a winning committee with respect to the profile p, then for any profile p' obtained from p by shifting a one position forward into a voter's preference while keeping all the other preferences unchanged, then a should still belong to a winning committee with respect to p'.

The *neutrality* property is also well known in the literature of social choice theory. It is a typical condition that is considered to ensure equal treatment of alternatives and requires the result of an election to change in compliance with renaming of alternatives. Let σ be a permutation of the set of candidates A and p be a preference profile. Let $\sigma(p) = (\sigma(p_1), \cdots, \sigma(p_n))$ be a new profile such that for any voter $i \in N$ and all $a, b \in A$, $\sigma(a)\sigma(\succ_i)\sigma(b) \Leftrightarrow a \succ_i b$. The property of neutrality has originally been defined as follows: For any committee size $k \leq m-1$, we have $F\Big(\sigma(p), k\Big) = \sigma\Big(F(p, k)\Big)$, where $\sigma\Big(F(p, k)\Big) = \Big\{\sigma(W) : W \in F(p, k)\Big\}$. In our setting, we can naturally adjust this property by making some restrictions on the set of possible permutations σ such that for every candidate $a \in A, j(a) = j(\sigma(a))$; that is, the permutations changing any candidate into a candidate with the same type. In this case, we can check that any diverse committee scoring rule satisfies this property. Without the above restriction, we cannot insure that the property is satisfied.

The *non-imposition* property is also well known in the literature of social choice theory and requires in its original form that any committee having the target size can be selected. It is clear that this property is failed by every diverse committee scoring rule since we impose in our selection process that the only committees that could be chosen are those that respect the diversity constraint or those that come as close as possible to the diversity constraint.

Note also that some properties defined by Lang and Skowron (2018) in order to select diverse committees in the multi-attribute setting can be reshaped to our model.[12] *House monotonicity*, for instance, is trivially satisfied by every diverse separable committee scoring rule. The house monotonicity property can be defined in our model as follows: If a committee W is a winning committee with respect to the triplet (p, k, q) and a committee W' is a winning committee with respect to the triplet (p, k', q) with $k' > k$, then, for any type $j \in [l]$, the number of candidates with type j in W' is greater than (or equal to) the number of candidates with type j in W. However, it seems to us that some other properties defined by Lang and Skowron (2018), such as the *non-reversal* property, are not meaningful in our framework. Literally, the non-reversal property requires that, if the

[12]Note that those properties are generalisations of axioms commonly considered in the political science literature in the context of apportionment (Balinski and Young, 1979, 2001).

diversity quota of a type j is greater than the diversity quota of another type j', then the proportion (or number) of the selected candidates of type j should be greater than (or equal to) the proportion of the selected candidates with type j'. This condition is no longer relevant in our framework since the diversity quota enforced for each class is the minimal number of candidates to be selected, but not the exact number of candidates to be selected from the class.[13]

5 Concluding comments

In this paper we considered the problem of selecting sets of candidates of a predefined size (committees) on the basis of the preferences of the voters. Specifically, we were interested in committee selection subject to diversity constraints. We assumed that the set of candidates is partitioned into classes (such as classes of men and women) according to a specific attribute (such as gender) and the goal is to select a committee that, on the one hand, has the highest possible score regarding a given excellence measure, but that, on the other hand, meets certain diversity specifications regarding the composition of the committee. We have shown in this paper how the well-known class of committee scoring rules can be extended and used for the selection process in this framework. This extension is called throughout the paper the class of diverse committee scoring rules. The next step was to present some new properties that are desirable in our framework. We considered adaptations of well-established properties from the literature dealing with the standard committee selection framework and, further, we introduced new axioms that are specific to the task of committee selection under diversity constraints. Our next goal was mainly to study the behavior of the class of diverse committee scoring rules by testing those properties.

A number of questions remain open for future work, but it seems to us that among all, two directions are most pressing. The first direction is to apply our framework to Condorcet-based rules. The class of diverse committee scoring rules that we presented in this paper are fundamentally different from the rules that are motivated by the Condorcet principle. It would thus be interesting to investigate how diversity constraints should be dealt with when the selection process is based on pairwise majority comparisons. The second direction would be to follow the work started by Bredereck et al. (2018) who introduced the concept of "price of diversity", which quantifies the "cost" of introducing diversity constraints in the committee selection setting. Basically, the committees that maximize the excellence score using a committee scoring rule are not necessarily diverse (with regard to the diversity constraint) and, then, we sometimes need to "pay a price" in order to select a diverse committee. Bredereck et al. (2018) express the price of diversity as the ratio between the score of the committee that would have won without considering the diversity constraint and the score of the committee that wins when the diversity constraint

[13] Similarly, it seems to us that *exactness and respect of quota* and *population monotonicity* defined by Lang and Skowron (2018) are not compatible with our setup for the same reasons.

is taken into account. We believe that this concept deserves more attention.

References

Aziz, H. (2019). A rule for committee selection with soft diversity constraints. *Group Decision and Negotiation*, 28(6):1193–1200.

Balinski, M. and Young, H. (1979). Criteria for proportional representation. *Operations Research*, 27(1):80–95.

Balinski, M. and Young, P. (2001). Fair representation. *2nd edition, Brookings institution press.*

Barberà, S. and Coelho, D. (2008). How to choose a non-controversial list with k names. *Social Choice and Welfare*, 31(1):79–96.

Bei, X., Liu, S., Poon, C. K., and Wang, H. (2022). Candidate selections with proportional fairness constraints. *Autonomous Agents and Multi-Agent Systems*, 36(1):1–32.

Brams, S. J. (1990). Constrained approval voting: A voting system to elect a governing board [with comment]. *Interfaces*, 20(5):67–80.

Bredereck, R., Faliszewski, P., Igarashi, A., Lackner, M., and Skowron, P. (2018). Multi-winner elections with diversity constraints. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32.

Bubboloni, D., Diss, M., and Gori, M. (2020). Extensions of the Simpson voting rule to the committee selection setting. *Public Choice*, 183(1):151–185.

Celis, L. E., Huang, L., and Vishnoi, N. K. (2017). Multiwinner voting with fairness constraints. *arXiv preprint*, https://arxiv.org/pdf/1710.10057.pdf.

Chamberlin, J. R. and Courant, P. N. (1983). Representative deliberations and representative decisions: Proportional representation and the Borda rule. *The American Political Science Review*, 77(3):718–733.

Diss, M. and Doghmi, A. (2016). Multi-winner scoring election methods: Condorcet consistency and paradoxes. *Public Choice*, 169(1): 97–116.

Diss, M., Kamwa, E., and Tlidi, A. (2020). On some k-scoring rules for committee elections: agreement and Condorcet Principle. *Revue d'Economie Politique*, 130(5):699–725.

Diss, M. and Mahajne, M. (2020). Social acceptability of Condorcet committees. *Mathematical Social Sciences*, 105:14–27.

Do, V., Atif, J., Lang, J., and Usunier, N. (2021). Online selection of diverse committees. *arXiv preprint*, https://arxiv.org/pdf/2105.09295.pdf.

Ehlers, L., Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2014). School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683.

Elkind, E., Faliszewski, P., Skowron, P., and Slinko, A. (2017). Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3):599–632.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2018). Multiwinner analogues of the Plurality rule: Axiomatic and algorithmic perspectives. *Social Choice and Welfare*, 51(3):513–550.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2019). Committee scoring rules: Axiomatic characterization and hierarchy. *ACM Transactions on Economics and Computation*, 7(1):1–39.

Gehrlein, W. V. (1985). The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10(3):199–209.

Goto, M., Kojima, F., Kurata, R., Tamura, A., and Yokoo, M. (2017). Designing matching mechanisms under general distributional constraints. *American Economic Journal: Microeconomics*, 9(2):226–62.

Ianovski, E. (2022). Electing a committee with dominance constraints. *Annals of Operations Research*, 318:985–1000.

Izsak, R., Talmon, N., and Woeginger, G. (2018). Committee selection with intraclass and interclass synergies. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32.

Kagita, V. R., Pujari, A. K., Padmanabhan, V., Aziz, H., and Kumar, V. (2021). Committee selection using attribute approvals. In: *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 683–691.

Kamada, Y. and Kojima, F. (2015). Efficient matching under distributional constraints: Theory and applications. *The American Economic Review*, 105(1):67–99.

Kamwa, E. (2017). On stable rules for selecting committees. *Journal of Mathematical Economics*, 70:36–44.

Kamwa, E. and Merlin, V. (2015). Scoring rules over subsets of alternatives: Consistency and paradoxes. *Journal of Mathematical Economics*, 61:130–138.

Kilgour, D. M. (2010). Approval balloting for multi-winner elections. In: *J. Laslier and M. R. Sanver (Eds.), Handbook on approval voting* (pp. 105-124), Heidelberg: Springer

Kilgour, D. M. and Marshall, E. (2012). Approval balloting for fixed-size committees. In: *Electoral Systems*, pp. 305–326. Springer.

Lang, J. and Skowron, P. (2018). Multi-attribute proportional representation. *Artificial Intelligence*, 263:74–106.

Lu, T. and Boutilier, C. (2011). Budgeted social choice: From consensus to personalized decision making. In: *Twenty-Second International Joint Conference on Artificial Intelligence*.

Ratliff, T. C. (2003). Some startling inconsistencies when electing committees. *Social Choice and Welfare*, 21(3):433–454.

Relia, K. (2021). Dire committee: Diversity and representation constraints in multiwinner elections. *arXiv preprint*, https://arxiv.org/pdf/2107.07356.pdf.

Skowron, P., Faliszewski, P., and Lang, J. (2016). Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216.

Skowron, P., Faliszewski, P., and Slinko, A. (2019). Axiomatic characterization of committee scoring rules. *Journal of Economic Theory*, 180:244–273.

Staring, M. (1986). Two paradoxes of committee elections. *Mathematics Magazine*, 59(3):158–159.

Thejaswi, S., Ordozgoiti, B., and Gionis, A. (2021). Diversity-aware k-median: Clustering with fair center representation. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 765–780. Springer.

Young, H. P. (1974). An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1):43–52.

Young, H. P. (1975). Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838.

Compliance with Ethical Standards

All authors declare that they have no conflicts of interest.