

Descriptive and Substantive Representation in Approval-Based Committee Voting

MSc Thesis (*Afstudeerscriptie*)

written by

Kirti Singh

under the supervision of **Prof. Dr. Ulle Endriss**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense: **Members of the Thesis Committee:**

July 9, 2025

Dr. Balder ten Cate (chair)

Prof. Dr. Ulle Endriss (supervisor)

Prof. Dr. Davide Grossi

Prof. Dr. Guido Schäfer



INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Approval-based committee voting (ABC voting) is a form of multi-winner voting in which all voters admit a set of candidates they approve of as their ballot. After all ballots are collected, a winning committee of size k is chosen. An important class of axioms in this context are the *proportionality axioms*. These capture the idea that the support of a candidate should be proportional to their presence in the winning committee. Investigating this class of axioms is part of the larger project of investigating fairness in multi-winner voting. This thesis can be regarded to be a part of this larger project as well.

In the context of political elections, representing voters based on their preferences for certain candidates is seen as a kind of *substantive* representation. This form is contrasted with *descriptive* representation. The latter emphasizes the demographic similarity of the representer to the representee, whereas the former emphasizes that the representer should work according to their voters' interests. The beginning of this thesis contains an argument for the pursuit of descriptive representation in many multi-winner voting contexts. After this, it will be argued that the kinds of proportionality and fairness that have been studied in the ABC context are preference-based, that is, substantive. This paves the way for the main contribution of this thesis: a formalisation and study of descriptive representation in the context of ABC voting by proposing an extended version of the ABC model.

The middle sections contain the definition of two classes of voting rules and their axiomatic analysis. For this, established ABC axioms are generalised to the ABCC setting and the two rules are compared on the basis of their performance with respect to these axioms and their computational properties. One of the conclusions is that both rules perform poorly with respect to proportionality axioms, suggesting incompatibility of extreme forms of descriptive and substantive representation. Furthermore, determining the winning committees for both classes of rules is intractable. In the final section, I try to understand this incompatibility by considering domain restrictions for which the rules are well-behaved. I finish by proposing an alternative class of voting rules with the aim of balancing substantive and descriptive representation.

Acknowledgements

First, I would like to thank my supervisor, Ulle, for guiding me through the thesis-process. After you suggested that *voter representation* had not been investigated systematically in multi-winner voting, you gave me a lot of freedom to explore this idea further, which I greatly enjoyed. Everything went quite smoothly, which would not have been possible without your support. Preparing for the biweekly meetings gave me a lot of structure and helped in distilling my ideas. I would also like to thank Ronald, for thinking about the complexity of some rules I had developed. Defending my thesis in front of Balder, Davide and Guido, my committee, was a welcome challenge. I would like to thank the committee members for reading my thesis and sharing their ideas and critiques with me.

My student life was a tale of two cities, taking place in Nijmegen and Amsterdam. I'm thankful for all the people who welcomed me when I came to the 'big city' of Nijmegen as a big-eyed boy from de Achterhoek. I'm also thankful for the friends I made during my philosophy and mathematics courses in Nijmegen. After my first course on logic, I realised that this might be a topic I could pursue for my master's. As time went on, this feeling grew stronger and I also realised that Nijmegen was not such a big city after all. This led me to apply for the Master of Logic. After moving around Amsterdam a couple of times, I finally landed in the *Logic House*. Living and studying in Amsterdam, but especially in the Logic House, has been a great experience. For this, I'm thankful for my housemates and the friends I made during the master, especially the computational social choice group from our cohort. Discussing the course material (and my thesis ideas!) was both helpful and a lot of fun. All in all, I'm thankful for all the friends I have made (and lost) along the way.

Lastly, and most importantly, I would like to thank my parents, Balwinder Singh and Ranjit Kaur, for wholeheartedly supporting me throughout my studies, even if you did not completely understand why I chose the subjects I did. Now, I have come to realise that a big part of the reason for this was that I never took the time to explain what I was doing. I am glad that this is not the case any more. I would also like to thank my little sister, Birta. There is no one in this world that understands me like you do.

Contents

Introduction	2
1 Preliminaries	9
1.1 ABC voting and basic definitions	9
1.2 Voting rules	10
1.3 Proportionality and representation	13
1.4 Apportionment methods	16
2 Separate electorates	19
2.1 Descriptive and substantive representation	19
2.2 Two perspectives on separate electorates	24
2.3 The model	25
2.4 Observing the two procedures	28
3 Axiomatic and computational properties	34
3.1 Basic axioms	35
3.2 Monotonicity	38
3.3 Consistency	44
3.4 Proportionality and representation	46
3.5 Strategic voting	49
3.6 Computational properties	51
3.7 Discussion	54
4 Escape routes	57
4.1 Conditions and (domain) restrictions	59
4.2 Community Phragmén-like rules	66
4.3 Discussion	70
Conclusion	72
Bibliography	77
A Implementation	82
A.1 RSP and NWP as algorithms	82
A.2 Results of the systematic generation of elections	83

Introduction

Nearly all modern democracies are forms of *representative democracies*, meaning that, instead of voters deciding on issues themselves they elect a *representative*, usually by voting for them in a parliamentary election. But what does it mean to be ‘represented’ in this context? Is it seeing the person you vote for end up in a lower house? Is it having someone in power who passes legislation in line with your personal preferences? Or is it someone who belongs to the same ‘group’ you belong to? Furthermore, *who* should be represented? All voters? As members of a certain group? Or as individuals? Consider the case in which voters with similar preferences have to elect two representatives. Suppose furthermore that the possible representatives share the same preferences and interests as their voters. In this case, would it be fairer to elect two people from the same group or one person from one group and one from another? Or, concretely, would it be fairer to elect two men or a man and a woman?

This thesis tries to answer questions related to the ones asked in the paragraph above. It is possible to approach these questions from at least two perspectives. A political scientist can analyse the institutional and political conditions that lead to the most representative outcomes. For this, one has to investigate what it means to be represented (Bird et al., 2010; Pettit, 2012; Ruedin, 2020). The other perspective is *social choice theory*, specifically, *computational social choice* (Brandt et al., 2016a). This is a discipline devoted to the study of situations in which groups of agents have to make collective decisions. These are analysed by employing formal tools from mathematics and computer science. The field that will be most important for this thesis is *multi-winner voting*. As the name suggests, this is the situation in which a group of voters cast their votes in order to elect multiple winners. Parliamentary elections can be seen as examples of this.

A lot of work has been done on representing voters in the context of multi-winner voting, with the most important notions of representation being *extended justified representation (EJR)*. This can be first found in the work of Aziz et al. (2017). Other notions of representation have been developed as well, by Sánchez-Fernández et al. (2017) and Peters and Skowron (2020). A property these notions have in common is that they represent voters based on their preferences and do not take into consideration the identity of the voters. This highlights two different perspectives on voting: *substantive representation* and *descriptive representation* (Ruedin, 2020). The former captures the idea that voters should be represented by people that work in their interests, no matter their identity. Whereas the latter emphasizes the similarity in demographic traits of the *representer* and the *representee*. The descriptive form of representation has not been interpreted in the context of multi-winner voting from the perspective of computational social choice. This is the main observation motivating this thesis.

This introduction contains a short overview of the field of computational social choice. Afterwards, we will consider established forms of proportionality in *approval based committee voting (ABC voting)*, a kind of multi-winner voting, and state the main research question. Lastly, there is an outline of this thesis.

Computational social choice

Computational social choice is a field of research that is indebted to two other disciplines: *social choice theory* (SCT) and *computer science*. Whenever we systematically analyse what a collective of individuals ought to decide in a given situation, we are doing SCT (Arrow et al., 2002; Suzumura, 2002). As Fishburn (1973) put it, it is the field that is “concerned with relationships between individuals’ preferences and social choice”. As mentioned before, a situation that fits this description is voting in national elections. *Voting theory* also plays a central role in SCT. For an exhaustive overview of computational social choice, the reader is referred to Brandt et al. (2016a). An overview of voting theory can be found in the same handbook (Zwicker, 2016).

To formally analyse voting, we model the situation as follows. First, there is a set of *voters* $N = \{1, 2, \dots, n\}$ that can vote on a set of *alternatives* $A = \{1, 2, \dots, m\}$. In the case of national elections, N corresponds to all eligible voters and A to politicians or political parties. Each voter submits a *ballot*. In classical voting theory, this is a linear order — called the *preference order* — over the full set of candidates. An example for $m = 3$ is: $a \succ b \succ c$. The ballot for voter i is denoted by $A(i)$. The ballots of all voters can be collected as a *profile* $A = (A(1), A(2), \dots, A(n))$, which we visualise as follows:

1 : $a \succ b \succ c$
2 : $b \succ c \succ a$
3 : $c \succ a \succ b$

$i : c_1 \succ c_2 \succ \dots \succ c_m$ is read as ‘ i submits $c_1 \succ c_2 \succ \dots \succ c_m$ as her ballot’. Essentially, a profile contains the ballots of each voter. Once the preferences of each voter have been submitted, the most important question that follows is: *who do we select as the winner?* or *what is the most appropriate outcome?*. This is done using a *voting rule* $\mathcal{R}$. These are functions that take profiles as input and return winning candidates. A straightforward rule is the *plurality rule* which returns the candidate that is ranked first by the most voters (Zwicker, 2016). Applying this rule to the profile above leads to a tie between a , b and c . Note that a , b and c are not only ranked first the same number of times, but they are ranked the same number of times on each position!

This example points towards different historical periods in the development of social choice theory. Roughly, there are three phases: the *pre-axiomatic*, *axiomatic* and *computational* phase. The first phase starts in antiquity, where Aristotle in ancient Greece, and Kautilya in ancient India investigated different forms of political and collective decision making in works titled *Politika* (*politics*) and *Arthashastra* (*science of polity*), respectively (Suzumura, 2002). It is important to note that in this time, the authors were not aware of the fact that they were doing SCT and their treatises dealt with general political and economic problems. The pre-axiomatic phase ended roughly around the beginning of the twentieth century.

The profile above was first used in an argument by Marquis de Condorcet (1743 – 1794). Hence, it is also known as *Condorcet’s paradox*. It captures the idea that it is impossible to always have majority support for an output. Another important figure in this time was Jean-Charles de Borda (1733 – 1799) who proposed an alternative to the plurality rule, after Condorcet pointed out the weak spot. The idea of Borda’s rule is that for each voter a candidate gets as many points as candidates ranked below her. This leads to a score for each candidate. Usually, the candidate with the highest score is elected as the winner. Note that during this period of the development of SCT, particular

rules were compared in different contexts. Condorcet investigated the plurality rule, whereas Borda argued for the superiority of his rule because it performed better for certain other examples (Brandt et al., 2016b).

The jump from *particularity* to *generality* marks the rupture between the pre-axiomatic and axiomatic phases. Instead of analysing voting rules, this period is marked by the formulation of general criteria voting rules should satisfy. These are known as *axioms*. Roughly speaking, demanding that voting rules satisfy certain axioms poses restrictions on our choice of rules. For instance, assuming *surjectivity* forces the voting rule to return each candidate at least once and excludes dictatorial rules, or, rules that output one candidate for every input. The investigation of different combinations of axioms is what characterises this period. This gave rise to *impossibility* and *characterisation* results. Arguably the most well known of these results is *Arrow's impossibility theorem* published by Kenneth J. Arrow (1951). Arrow showed that a voting rule satisfying two seemingly basic and fair axioms must be a dictatorship. A full statement and proof of Arrow's theorem can be found in the overview of Campbell and Kelly (2002).

Another important impossibility theorem, which will help in understanding the turn from the axiomatic to the computational phase, is the *Gibbard-Satterthwaite Theorem* (Gibbard, 1973; Satterthwaite, 1975). They found that any voting rule that is *resolute*, *nonimposed* and *strategyproof* is a dictatorship. Resoluteness forces the voting rule to output one winner. Nonimposedness is the same as surjectivity and strategyproof captures the idea that voters should not be able to vote strategically, i.e., submit a non-truthful ballot, in such a way that their satisfaction with the output increases. A way of relativising the negativity of this impossibility theorem is by analysing the *computational complexity* of voter manipulation. Here, we ask the question: how hard is it for a voter to strategise in such a way that she increases her satisfaction? If it is computationally *hard* to do this, it is unlikely that a voter will actually be able to strategise successfully.

Computational results like these mark the *computational turn* from social choice theory to computational social choice that occurred in the beginning of the twenty-first century. Next to analysing the computational complexity of voter manipulation, another important question one can ask is how hard it is to determine the winner of a given voting rule. Voting rules that seem *normatively appealing* and satisfy many desirable properties become less useful if it is not possible to compute their outcome in a reasonable amount of time. Next to this, the application of techniques from automated reasoning, such as using SAT-solvers, is a big part of computational social choice. One example is the systematic generation of impossibility results by Geist and Endriss (2011).

The standard voting theory setting, that has persisted throughout the three phases, demands each voter to rank all candidates linearly and outputs one winner only. Both of these elements can be generalised and there are practical reasons to do so. Ranking all candidates linearly is a cognitively demanding task. In the 2021 Dutch parliamentary elections, for instance, there were 37 participating parties. One of them being the *Feestpartij (the Party Party)* (Kiesraad, 2021). Ranking 37 objects in a voting booth is a lot to ask. One can weaken what we ask from the voters to, for instance, a partial ranking or an *approval ballot*. Furthermore, instead of outputting one winner, it is also possible to vote on multiple winners. The combination of this leads to the main social choice context we will be delving into for this thesis: *approval ballot voting*.

In approval ballot voting, each voter i admits a subset of candidates they approve of, i.e., a subset $A(i) \subseteq C$. The goal is to find a winner set W , which is also known as the *winning coalition*, of size

k . The main arguments for the adoption of this system are its generality and its applicability, and the cognitive simplicity of the ballots. It also lets voters provide more information than single ballots, which increases voter engagement and participation (Brams and Fishburn, 2010). Examples of when we could use approval ballot voting are: parliamentary elections, going for dinners with friends and ordering dishes, selecting a jury that has to deliberate a case and locating public facilities. For more applications the reader is referred to the overviews by Lackner and Skowron (2023) and Laslier and Sanver (2010).

Proportionality

A big part of the work on approval-ballot voting is about *proportionality* and questions about fairness. There are multiple winners, so, it is possible to proportionally represent certain parts of the electorate with certain parts of the final committee. To illustrate this, we consider an example. Suppose that there are $n = 100$ voters, $m = 7$ candidates and we want to elect $k = 2$ winners. This is one illustration of a profile:

$$\begin{aligned} 51 &\times \{l, b\} \\ 18 &\times \{c, s\} \\ 16 &\times \{g, s\} \\ 15 &\times \{p, s\} \end{aligned}$$

$j \times X$ denotes ‘ j voters approve of X ’. If we elect the candidates that get the most approvals, $\{l, b\}$ is the final winning committee. Electing $\{l, s\}$ or $\{b, s\}$, however, is more *diverse*, as now everyone is represented by at least one voter. It is also a better image of the voter preferences. Fairness and proportionality can be formalised as axioms. With the most well-known notion of fairness being *extended justified representation* developed first Aziz et al. (2017). This notion of proportionality captures the idea that a group of voters that is ‘big enough’, i.e., of size at least $\frac{n}{k}$, and that is ‘cohesive enough’ in terms of preferences, which means they share a certain amount of candidates, must be granted a certain amount of representatives.

This thesis also deals with proportionality. The distinguishing factor of this thesis is that it deals with another kind of proportionality. Note that EJR defines ‘cohesiveness’ in terms of voters’ preferences and not in terms of demographic traits a voter has. Representation based on interests and preferences is called *substantive representation*, whereas representation by ‘somebody from your own group’ is called *descriptive representation* (Ruedin, 2020). The raison d’être of this thesis is introducing descriptive representation to the context of multi-winner voting and comparing it to established notions of proportionality that represent voters based on their approvals. The main question is:

Is it possible to represent voters in the ABC model? If so, in how far is it possible to combine this form of representation with substantive forms of representation, in particular, extended justified representation and perfect representation?

In Chapter 2 we will argue why descriptive representation should be pursued in the context of parliamentary elections. Next to this context, there are other cases in which the identity of a voter plays a substantial role. I list some of these examples in the remainder of this section.

Example 1 (Dinner with allergic friends). Suppose that you are having dinner with 5 friends and one of your friends is allergic to gluten and lactose. You want to select 2 starters to share. These are the approval ballots:

$$\begin{aligned}
&2 \times \{\text{pizza, hamburger, sushi}\} \\
&2 \times \{\text{pizza, hamburger, cheese platter}\} \\
&1 \times \{\text{sushi}\}
\end{aligned}$$

The one person approving only of the sushi is your allergic friend. In this context, it makes sense to make the allergic friend order something separately. First of all, the dish she prefers, sushi, is approved by two other friends. Second of all, most importantly, if she does not know for certain that she will get a dish, she might end up going home on an empty stomach. So, in this context, there is a sensible argument to be made to ‘fix’ one dish for the allergic friend. The rest of the dishes are selected as a collective.

Example 2 (Jury for police case). Suppose that you want to elect a committee of 12 people that will function as a jury during a case. Suppose furthermore that you live in a country with multiple minorities and that these minorities are involved in this case. The goal is to form a committee that is an image of the population in terms of background. We want to choose 12 jury-members. Suppose that this is the profile:

$$\begin{aligned}
&90 \times \{a_1, a_2, \dots, a_{100}\} \\
&5 \times \{a_1, b_1, b_2, \dots, b_5\} \\
&3 \times \{a_1, c_1, c_2, \dots, c_5\} \\
&2 \times \{a_1, d_1, d_2, \dots, d_4\}
\end{aligned}$$

a_1 belongs to group a , but she is a general representative who is liked by everyone. In general, all x_i belong to community x for $x \in \{a, b, c, d\}$. Here, it also makes to sense to ‘reserve’ certain parts of the committee for people belonging to different minorities, to ensure representation. If we go off preferences, choosing a_1 would ensure that ‘everyone is happy’, but choosing $\{a_1, a_2, \dots, a_{12}\}$ is not representative. An alternative approach is fixing that each group gets a fixed part of the jury.

Example 3 (Public toilets). Amsterdam is known for its *plaskrullen* (*pee curls*). These are open-air pissoirs scattered throughout the city’s centre, built in the beginning of the twentieth century. The municipality has decided to not repair these public toilets if they are broken. Instead, they will replace them by toilets that are accessible by people of all genders (Haye, 2025). The reason behind there only being men’s public toilets is that the toilets were introduced in times when women did not participate in municipal governing on a large scale. This story can be translated to an ABCC problem. Suppose that there are 100 people approving of three types of toilets: open air *plaskrullen*, closed off *flush toilets* and closed off *squatting toilets*. The city wants to build 50 new toilets This is the profile:

$$\begin{aligned}
&46 \times \{\text{plaskrullen}\} \\
&5 \times \{\text{plaskrullen, squatting toilets}\} \\
&49 \times \{\text{flush toilets}\}
\end{aligned}$$

Suppose furthermore that the 51 voters approving of plaskrullen are men and the 49 approving of the flush toilets are women. Electing what the majority votes for would lead to us deciding for plaskrullen only. We consider two solutions for this neglect of the 49 votes. A straightforward solution is to use a proportional or diverse voting rule. Another way of tackling this problem is assigning 25 toilets to the women and 25 toilets to the men. Afterwards, they can decide what kind of toilets they want.

Example 4 (Medicine testing). Many medicines have different (negative) effects on different groups of people; be it different ethnic groups or different genders. One way of setting up a medical research is by having diverse trial groups test the medicine. In a typical setting, a centralised research collectively decides whether a medicine is fit for large-scale prescription. In many cases, these centralised teams skew towards certain demographic traits. Instead of having a centralised decision, a way of deciding on the safety of a medicine could be based on belonging to high-risk groups. Each high-risk group would have the ability to approve of certain medicines. This is another instance in which the demographic traits do matter.

Related work

As this thesis delves into proportionality, the whole surveying section on proportionality from Lackner and Skowron (2023) can be seen as related work. In particular, their section on approval voting with external properties is relevant. Note, however, that this is proportionality that I deemed to be of the substantive and not of the descriptive kind. Related papers from the ABC literature are papers on proportionality that include some kind of philosophical reflection on their definitions and framework. This includes the work by Aziz et al. (2017), where they first define EJR and the paper by Peters and Skowron (2020). Sánchez-Fernández et al. (2017) criticise the framework of Aziz et al. (2017) on the basis of another intuitive definition of proportionality. Their work shows similarity with Monroe (1995) who defines a voting rule based on considerations about fairness. These considerations are extended and further contextualised in this thesis.

Representing legislators based on one extra attribute is done in many places in the world, most notably in India, Pakistan and Lebanon. Work on electing candidates that also belong to a certain group, such as a party or community, is called *bi-apportionment* or *double proportionality*. An overview of this can be found in Pukelsheim (2017). The *apportionment* context is different from the approval voting setting. It can be seen as a restricted form, in which there are parties with candidates that voters vote for as a whole and not for individual candidates. This makes it related, but essentially distinct. The apportionment setting in which voters can approve of more than two attributes is called *multi-attribute apportionment voting* and it is studied by Lang and Skowron (2018).

The first formalisation of multi-attribute *approval* voting is due to Kagita et al. (2021). This model will be discussed extensively in Chapter 2. The idea is that voters do not only approve of candidates, they also approve of attributes this candidate might possess. For instance, in the context of parliamentary elections a voter could approve of the region a candidate is from, the gender and the education level. This could lead to an approval ballot of the following form $\{(city, woman, BA)\}$. The main problem with their formal model is that the proportionality axioms they propose are too strong to be satisfied by any of the voting rules they define.

Lastly, it is important to mention work from the field of political science on representation in politics, as it is an important source of inspiration and an important application. The surveying work by Ruedin (2020) provides an excellent overview of both theoretical and experimental work done on representing minorities in representative democracies. Mansbridge (1999) provides a discussion of whether or not descriptive representation should be pursued from the perspective of the United States. Pettit (2012) is an important theoretical source that provides an overview of different kinds of public representation. Further sources which focus on different democratic systems with built-in

representation guarantees are Lijphart (1969) and Rai (2005) focussing on South Asia, and Tan and Preece (2022) with a focus on Asia Pacific.

Outline

Chapter 1: Preliminaries This chapter contains the basic requirements to understand the main contribution of this thesis. It contains a formal exposition of the *approval based committee voting* model, voting rules from this system, the main proportionality axioms that have been mentioned in this introduction and an overview of some relevant apportionment methods.

Chapter 2: Separate electorates The beginning of this chapter is devoted to motivating descriptive representation. It is an outlook on the main arguments against a focus on substantive representation only, and contains arguments in favour of a form of descriptive representation. It also contains the exposition of a modified version of the ABC model, which is designed to enforce a form of descriptive representation on the design level. This will be dubbed the *ABC model with communities*, or, the *ABCC model*. It also contains the motivation for and definitions of two ABCC voting procedures. The last part of this chapter is devoted to an analysis of these two voting procedures in practice. The results and discussion in this chapter are of a ‘pre-axiomatic’ nature.

Chapter 3: Axiomatic and computational properties This chapter contains content that typically falls into the axiomatic and computational categories. Each section contains a different class of axioms that are generalisations of the most important and well-known ABC axioms to the context of ABCC voting. The two voting procedures are compared on an axiomatic level. Lastly, the section contains the definition of basic decision problems for the context of ABCC voting. The final verdict will be that the two defined rules do not perform as desired. In particular, with respect to different kinds of substantive representation.

Chapter 4: Escape routes Since the two defined voting procedures do not perform as desired, we have to find some ‘escape routes’ for this problem. This section is a discussion of possible directions one could further investigate. Next to this, this section also contains an evaluation of the *raison d’être* of this model, building on the second chapter. The three escape routes that will be presented are: (i) domain restrictions — *one should only use this model in specific cases* —, (ii) new classes of voting rules — *the defined rules are too naive and simplistic* — and (iii) axioms specific to the ABCC model — *generalised ABC axioms do not paint a good picture of ABCC voting rules*.

Chapter 1

Preliminaries

1.1 ABC voting and basic definitions

In this section, the model I will be contributing to will be presented. I start with a motivating toy election:

Example 1.1. Suppose that a country with 100 people is choosing a prime minister and a president. Instead of voting for these two positions separately, the voters *approve* some candidates and then a committee of two winners is selected. It is important to remark that the voters themselves are allowed to decide *how many* candidates they approve of; they can approve exactly two, none, all or any other number of candidates. After two winners are selected, there is a bureaucratic process that decides who the president and who the prime minister will be. Each candidate is identified with one political ideology. Concretely, the people can choose from these candidates:

$$\{\underline{b}ig-tent, \underline{c}ommunist, \underline{g}reen, \underline{l}iberal, \underline{p}opulist, \underline{s}ocialist\}$$

This is the result after the election:

$$\begin{aligned} 51 \times \{l, b\} \\ 18 \times \{c, s\} \\ 16 \times \{g, s\} \\ 15 \times \{p, s\} \end{aligned}$$

The most straightforward way of picking two winners, is selecting the candidates with the most votes. If we do this, $\{l, b\}$ is the set containing the winners. $\diamond$

The winner set from the example above immediately raises a question: is this a *fair* outcome? Before it is possible to answer this question, we need more machinery to define fairness in this context. Hence, I now define the general form of this kind of social choice setting, which is known as *approval-based committee voting* or *ABC voting*. The main source for my exposition of this setting is the surveying work by Lackner and Skowron (2023). In the ABC setting there is a set $N = \{1, \dots, n\}$ of n voters and a set $C = \{c_1, c_2, \dots, c_m\}$ of candidates. Throughout this thesis, n refers to the cardinality of N , $|N|$, and m refers to $|C|$; as is the convention in the literature on ABC voting. Each voter $i \in N$ submits an *approval ballot* $A(i) \subseteq C$ containing all the candidates i ‘approves’ of. The profile $A = (A(1), A(2), \dots, A(n))$ is a vector containing the approval ballot of every voter. It is possible to

formalise it is a function $A: N \rightarrow \wp(C)$, but we omit this here. After receiving all the ballots, we want to select a *winner set* or *winning committee* $W \subseteq C$. The convention is to use k to refer to the cardinality of W ; $|W| = k$. An election is a tuple $E = (A, k)$ where A is the profile and $k \leq |C|$ the desired size of the winner set. Furthermore, let $N(C')$ denote the set of voters approving of *all* candidates in set $C' \subseteq C$, i.e., $N(C') = \{i \in N \mid c \in A(i) \text{ for all } c \in C'\}$. For brevity, I will write $N(c)$ instead of $N(\{c\})$.

It is now possible to formalise the example from above. Since this example will reappear more often in this preliminaries, I will refer to it as ‘the running example’.

Example 1.2 (Continued). Let $n = 100$, $C = \{b, c, g, l, p, s\}$ ($m = 6$) and $k = 2$. The profile is defined as follows: $A(i) = \{l, b\}$ for $i \in [1, 51]$, $A(i) = \{c, s\}$ for $i \in [52, 69]$, $A(i) = \{g, s\}$ for $i \in [70, 85]$ and $A(i) = \{p, s\}$ for $i \in [86, 100]$. The election instance in this case is $E = (A, 2)$. $\diamond$

1.2 Voting rules

The standard way of aggregating an election E into a winner set W is by using an *ABC voting rule* $\mathcal{R}$ that maps elections E into subsets of $\binom{C}{k}$. Here $\binom{C}{k}$ is the set containing all subsets X of C such that $|X| = k$. The literature on multi-winner voting and ABC voting contains many examples of voting rules, each with their own characteristics and use cases. For a survey I refer the reader to Chapter 3 of Lackner and Skowron’s (2023) book on ABC voting.

This section contains some of the most well-known and important voting rules. I start by presenting the most straightforward voting rule, which is called the *approval voting rule* or *AV-rule*. This is a formalisation of the ‘rule’ used informally in the running example. To define AV, we have to define its *score function*:

$$\text{score}_{\text{AV}}(A, c) = |N(c)| \text{ and } \text{score}_{\text{AV}}(A, W) = \sum_{c \in W} |N(c)|$$

Voting rule 1 (AV). The AV-rule, or $\mathcal{R}_{\text{AV}}$, maximises $\text{score}_{\text{AV}}(A, W)$. In other words, $\mathcal{R}_{\text{AV}}$ selects the k candidates which are approved by the most voters.

A rule that emphasizes diversity is *Approval Chamberlin-Courant*. This rule tries to maximise the number of voters with at least one candidate in the winner set. It can also be written as a rule that maximises a score function:

Voting rule 2 (CC). The CC rule, or $\mathcal{R}_{\text{CC}}$, outputs all committees W that maximise $\text{score}_{\text{CC}}(A, W) = |\{i \in N \mid A(i) \cap W \neq \emptyset\}|$.

This voting rule is named after John R. Chamberlin and Paul N. Courant, who presented it in a paper in (1983). The first formulation of this rule is due to Thorvald N. Thiele (1895).

The two presented rules belong to the general class of *Thiele methods* that try to maximise the *satisfaction* of voters. Here, satisfaction is defined as a function w , which is also called the *satisfaction function*. This function is determined only by the number of approved candidates each voter has in the committee W . In the case of AV it is $w(x) = x$ and in the case of CC it is $w(x) = \min(1, x)$. To understand the concept of Thiele methods, we first define them generally:

Definition 1.1 (Thiele methods). A Thiele method, or w -Thiele, is parametrised by a non-decreasing function $w: \mathbb{N} \rightarrow \mathbb{R}$ with $w(0) = 0$. w -Thiele maximises $\text{score}_w(A, W) = \sum_{i \in N} w(|W \cap A(i)|)$.

AV, as a Thiele method, maximises $\text{score}_w(A, W) = \sum_{i \in N} |W \cap A(i)|$, which means we maximise the number of votes W has. CC, maximises $\text{score}_w(A, W) = \sum_{i \in N} \min(1, |W \cap A(i)|)$. It is easy to see that this score is maximised if as many voters as possible have exactly one candidate they approve of in W .

In terms of satisfaction, we see that for AV the satisfaction of a voter increases linearly to the number of candidates in the final winner set; $w(x) = x$. In the case of CC, the satisfaction is maximised if there is just one candidate in the final winner set; $w(x) = \min(1, x)$. Put differently, AV can be seen as the Thiele method that is least diverse. The satisfaction function underlying AV assumes that voters' satisfaction keeps on increasing as we increase the number of candidates that represent them. CC is the most diverse Thiele method, since voters' satisfaction halts whenever they have one approved candidate in W .

A rule that is between AV and CC in terms of diversity is *proportional approval voting (PAV)*. The satisfaction function of this rule captures the idea of *diminishing returns*, i.e., the satisfaction of a voter does increase, but it does not increase linearly, instead it increases logarithmically. Formally, the satisfaction function of PAV is given by the Harmonic series, i.e., $h(x) = \sum_{j=1}^x \frac{1}{j}$. This leads to the following voting rule:

Voting rule 3 (PAV). PAV is the w -Thiele rule with $w(x) = h(x)$. In other words., it returns all committees that maximise $\text{score}_{\text{PAV}}(A, W) = \sum_{i \in N} h(|W \cap A(i)|)$

To show the working of these rules, I apply them to the running example.

Example 1.3. Again, we consider:

$$\begin{aligned} &51 \times \{l, b\} \\ &18 \times \{c, s\} \\ &16 \times \{g, s\} \\ &15 \times \{p, s\} \end{aligned}$$

Since $N(l) = N(b) = 51$ and for any other $c \in C$ we have $N(c) < 51$, it is the case that $\mathcal{R}_{\text{AV}}(E, 2) = \{l, b\}$. The following sets maximise the number of voters that are represented by at least one candidate: $\{\{l, s\}, \{b, s\}\}$. The total number of represented voters is equal to n , hence this is the output of $\mathcal{R}_{\text{CC}}$. $\mathcal{R}_{\text{PAV}}$ has the same output as $\mathcal{R}_{\text{CC}}$ in this case. The biggest 'increase' of satisfaction for the biggest number of voters occurs when we add l or b to W . The same increase of satisfaction occurs for 49 people whenever we add s . So, $\text{score}_{\text{PAV}}(\{l, s\}, 2) = \text{score}_{\text{PAV}}(\{b, s\}, 2) = 51 \cdot 1 + 49 \cdot 1$. Changing s to b decreases 1 to $\frac{1}{2}$ and changing it to anything else decreases 49 to 18, 16 or 15. $\diamond$

The remainder of this section is devoted to a presentation of more voting rules. Some of these do not belong to the class of Thiele methods or scoring rules. The first such rule we consider is *Monroe's rule*. This rule is similar to CC, but the main difference is that each winning candidate can represent at most $\frac{1}{k}$ of the voters. First, we define a class of functions that we call the *Monroe assignments*. This is a function $\varphi: N \rightarrow W$ such that for all $c \in W$ it is the case that $\lfloor \frac{n}{k} \rfloor \leq |\varphi^{-1}(c)| \leq \lceil \frac{n}{k} \rceil$. We regard each $\varphi(i)$ to be a representative of i . A Monroe assignment ensures that each candidate represents the same amount of people. Lastly, define $\Phi(W)$ as the set of all possible Monroe assignments for W .

Voting rule 4 (Monroe). The Monroe rule, or $\mathcal{R}_{\text{Monroe}}$ returns all committees with a maximum Monroe score, i.e., $\text{score}_{\text{Monroe}}(A, W) = \max_{\varphi \in \Phi(W)} |\{i \in N \mid \varphi(i) \in A(i)\}|$

This rule was first developed by Burt L. Monroe in 1995. For details, the reader is referred to the paper by Monroe (1995).

The following two rules are *non-standard ABC rules*. An ABC rule that is non-standard, does not coincide with AV for $k = 1$. To consider the first rule, we first define the *Hamming distance* between two sets X and Y as $d_{\text{Hamming}}(X, Y) = |X \setminus Y| + |Y \setminus X|$.

Voting rule 5 (Minimax approval voting (MAV)). MAV minimises the largest Hamming distance among all voters, i.e., $\mathcal{R}_{\text{MAV}}$ minimises $\max_{i \in N} d_{\text{Hamming}}(A(i), W)$

The second non-standard rule is defined as follows:

Voting rule 6 (Satisfaction approval voting (SAV)). SAV returns all committees that maximise $\text{score}_{\text{SAV}}(A, W) = \sum_{i \in N} \frac{|W \cap A(i)|}{|A(i)|}$

The last class of rules are the so-called *Phragmén rules*. This class of rules was first developed by Lars Edvard Phragmén (1863-1937). It is based on cost sharing, which means that each voter has a certain budget and that selecting a candidate comes with a certain cost. For more details on this class of rules, see the work by Brill et al. (2023). I first present a sequential method, which comes in two formulations.

Voting rule 7 (Sequential Phragmén). seq-Phragmén , or $\mathcal{R}_{\text{seqPhr}}$, can be formulated in two ways: as a continuous or as a round-based process.

Continuous formulation We start with an empty committee $W = \emptyset$. Each voter $i \in N$ starts with a budget $b_i = 0$. As time increases, the budget of each voter increases.

Candidates c can be ‘bought’ for a cost (or price) of $p_c = 1$. If a group of voters V has enough budget for a candidate, i.e., $\sum_{i \in V} b_i = p_c$ the group can collectively add it to W . After a candidate is added, the budget of all participating voters, $i \in V$, is reset to 0.

If $|W| = k$, W is returned.

Round-based formulation Start with an empty committee $W = \emptyset$ and iteratively add candidates that minimise the new maximum voter load. There are $r \leq k$ rounds. Let $y_i^{(j)}$ denote the cost a voter contributes after round j . Before we start, each voter has a contribution of 0: $y_i^{(0)} = 0$ for all $i \in N$.

Let $c_1, \dots, c_{r-1}$ be the candidates that have already been added to the committee in rounds 1 to $r - 1$. For each remaining candidate c , compute the maximum load that would arise from adding c :

$$\ell_r(c) = \frac{1 + \sum_{i \in N(c)} y_i^{(r)}}{|N(c)|}$$

Choose the candidate that keeps the maximum load as low as possible, i.e., we choose c_r such that $c_r = \arg \min_{c \in C \setminus \{c_1, \dots, c_{r-1}\}} \ell_r(c)$.

After choosing c_r , redefine the voter loads as follows: if $i \in N(c_r)$ $y_i^{(r)} = \ell_r(c_r)$ otherwise it is equal to the load from the previous round, $y_i^{(r-1)}$.

Return $W = \{c_1, \dots, c_k\}$.

The last rule we consider also is a round-based rule, which was first defined in the work by Peters and Skowron (2020).

Voting rule 8 (Method of Equal Shares (MES)). This rule is divided into two phases. Let $b_r(i)$ denote the budget of voter i in round r . Set $b_0(i) = \frac{k}{n}$ for all $i \in N$. Putting a candidate in the winner set

comes with a cost of 1. In round $r + 1$ consider the set of candidates that have not been picked, whose supporters are able to pay for them, i.e., all c such that $\sum_{i \in N(c)} b_r(i) \geq 1$. Add all such c to C_r .

If $C_r \neq \emptyset$, we find the minimal value $\rho(c)$ such that $\sum_{i \in N(c)} \min(\rho(c), b_r(i)) = 1$. We pick c such that $\rho(c)$ is minimised. b_{r+1} is defined as $b_r(i) - \rho(c)$ if $c \in A(i)$. Otherwise, $b_r(i) \geq \rho(c)$, $b_{r+1}(i) = 0$ if $c \in A(i)$ and $b_r(i) < \rho(c)$ and if $c \notin A(i)$ $b_{r+1}(i) = b_r(i)$. This captures the idea that voters who approve of c either pay $\rho(c)$ or their remaining budgets.

If $C_r = \emptyset$, we move on to the second phase. It might be the case that the W we have found is smaller than k . In this case, we have to add $k - |W|$ candidates to W . There is no standardised way to do this.

The main reason for presenting these rules that are difficult to grasp at a first reading is that they satisfy many desirable properties. One of the properties satisfied fall under the umbrella of proportionality and representation. Lastly, these rules are applied to the running example.

Example 1.4. First, we consider $\mathcal{R}_{seqPhr}$. At $t = \frac{1}{51}$, there are 51 voters in $N(\{l, b\})$ with a budget of $\frac{1}{51}$. Which means they can select l or b . We break ties in favour of l . Just shortly after, at $t = \frac{1}{49}$ the 49 voters in $N(s)$ can elect s . Hence, $W = \{l, s\}$. If we broke ties the other way around, $W = \{b, s\}$.

Now, we consider the MES. The initial budget of each voter is $\frac{1}{50}$. At round 1, there is a group that can pay for a candidate: $\sum_{i \in N(\{l, b\})} b_r(i) = \frac{51}{50}$. There are no other groups that can afford candidates. So, $C_r = \{l, s\}$. The minimum value each member of $N(\{l, b\})$ has to pay is $\frac{1}{51}$ to buy l or b . We break ties in favour of l and everyone in $N(\{l, b\})$ is left with a budget of $\frac{1}{50} - \frac{1}{51} = \frac{1}{2550}$. We now enter the second phase, since $|W| < 2$.

We use seq-Phragmén, where the initial budgets for the members in $N(\{l, b\})$ are $\frac{1}{2550}$ and for the rest the budgets are $\frac{1}{50}$. As time increases, their budgets increase as well. At $t = \frac{1}{49} - \frac{1}{50}$ there are 49 voters that can afford s . Hence, s is elected and $W = \{l, s\}$. $\diamond$

1.3 Proportionality and representation

A basic desire in the setting of ABC voting is fairness and the ability to represent voters. Therefore, *proportionality* has been studied in the literature extensively, with a highlight being the work of Aziz et al. (2017) on *(extended) justified representation (EJR)*. As mentioned in the introduction, EJR grants members of groups that are ‘cohesive and big enough’ some form of representation. Hence, to define EJR, we have to formalise the notion of cohesiveness. To do this, we define *ℓ -cohesive groups*; this is a group whose members have at least ℓ approved candidates in common and it satisfies a size constraint. The definition of EJR is from the work by Aziz et al. (2017), but the formulation is due to Lackner and Skowron (2023).

Definition 1.2. For $\ell \geq 1$, a group $V \subseteq N$ is *ℓ -cohesive* if:

1. $|V| \geq \ell \cdot \frac{n}{k}$.
2. $|\bigcap_{i \in V} A(i)| \geq \ell$.

We can now define extended justified representation:

Axiom 1. An ABC rule $\mathcal{R}$ satisfies *extended justified representation* if for each election instance $E = (A, k)$, each winning committee $W \in \mathcal{R}(E)$ and each ℓ -cohesive group of voters V there exists a voter $i \in V$ with at least ℓ representatives in W ; $|A(i) \cap W| \geq \ell$.

If we replace ‘each ℓ -cohesive group’ by ‘each 1-cohesive group’, we obtain the definition *justified representation* or JR. We investigate this axiom in the context of the running example.

Example 1.5 (Running example). Recall that $n = 100$, $m = 6$ and the profile is given by:

$$\begin{aligned} &51 \times \{l, b\} \\ &18 \times \{c, s\} \\ &16 \times \{g, s\} \\ &15 \times \{p, s\} \end{aligned}$$

If $k = 2$, $N(\{l, b\})$ is the only 1-coalition. Since $|N(\{l, b\})| = 51$ which is greater than $1 \cdot \frac{n}{k} = 50$. There is no bigger coalition, since this would require 100 members. To satisfy EJR, we need at least one voter from $N(\{l, b\})$ to be represented. This is the case for all rules presented in the previous section. $\diamond$

Note that the group of 49 voters does not satisfy the conditions for EJR, since $49 < \ell \cdot \frac{100}{2}$ for any $\ell \in \mathbb{N}_{\geq 1}$. Also, recall that $\mathcal{R}_{CC}$ and $\mathcal{R}_{PAV}$ output $\{\{l, s\}, \{b, s\}\}$ as the winner. These winning sets are ‘more proportional’ than EJR forces us to be. ‘More’ and ‘less’ proportional can be quantified and investigated by defining the *proportionality degree* (Lackner and Skowron, 2020b).

Definition 1.3 (Proportionality degree). Fix a function $f : \mathbb{N} \rightarrow \mathbb{R}$. An ABC rule $\mathcal{R}$ has a *proportionality degree* of f if for each $E = (A, k)$, each winning committee $W \in \mathcal{R}(E)$ and each ℓ -cohesive group V , the average number of representatives that voters from V get in W is at least $f(\ell)$, or:

$$\frac{1}{|V|} \cdot \sum_{i \in V} |A(i) \cap W| \geq f(\ell)$$

The proportionality degree is a quantitative method of assessing how fair a rule treats cohesive groups. As expected, AV is not a fair rule with respect to this metric. We will also see that many times f is a constant function, meaning that the number of representatives a ℓ -cohesive group gets is constant. One example of this is the following result. This result was first proven by Skowron (2021), but the following proof is original.

Proposition 1.1. $\mathcal{R}_{AV}$ has a proportionality degree of 0.

Proof. Fix $\ell \in \mathbb{N}$. Set $k = 2\ell + 1$, $n = k$ and $m = 2k$. Consider:

$$\begin{aligned} &\ell + 1 \times \{c_1, \dots, c_k\} \\ &\ell \times \{c_{k+1}, \dots, c_{2k}\} \end{aligned}$$

Note that the second community deserves ℓ candidates, as they are an ℓ -cohesive group. They get 0, since $W = \{c_1, \dots, c_k\}$. $\square$

Ideally, we want each member of a ℓ -cohesive group to be represented by ℓ candidates. Put differently, we want a voting rule that has a proportionality degree of $f(\ell) = \ell$. Unfortunately, Aziz et al. (2018) show that this is not possible.

Proposition 1.2 (Aziz et al., 2018). *There is no rule with proportionality degree $f(\ell) = \ell$.*

Proof. Let $n = 12$, $m = 4$ and $k = 3$:

1 : {a, d}	2 : {a}	3 : {a}
4 : {a, b}	5 : {b}	6 : {b}
7 : {b, c}	8 : {c}	9 : {c}
10 : {c, d}	11 : {d}	12 : {d}

There are four 1-cohesive groups: $\{1, 2, 3, 4\}$, $\{4, 5, 6, 7\}$, $\{7, 8, 9, 10\}$, and $\{1, 10, 11, 12\}$ for candidates a , b , c and d respectively. But $|W| = 3$, so, there always is a 1-cohesive group of voters that gets no representative. This means that they have a proportionality degree of less than $\ell = 1$. $\square$

The following two propositions show that it is possible to have a rule that is ‘more’ proportional than EJR. The first result is due to Sánchez-Fernández et al. (2017) and the second is due to Aziz et al. (2017). Proofs of these results are quite technical and can be found in these sources.

Proposition 1.3 (Sánchez-Fernández et al., 2017). *Any rule that satisfies EJR has a proportionality degree of at least $\frac{\ell-1}{2}$.*

The other direction does not hold. So there exist rules with proportionality degree of at least $\frac{\ell-1}{2}$ that do not satisfy EJR. Note that any rule satisfying justified representation, has a proportionality degree of ≤ 1 . This follows from the definition, as JR guarantees that each ℓ -cohesive group has at least one member that gets one representative. In the best case each member of the ℓ -cohesive group gets one representative. In the worst case, exactly one member gets a representative and the rest gets nothing. In this case, the proportionality degree is lower than 1.

Theorem 1.4 (Aziz et al., 2017). *PAV has a proportionality degree of $\ell - 1$. It also satisfies EJR.*

It is important to note that the proportionality degree can be quite low, while a rule has diverse outputs. This is the case because the proportionality degree caters to cohesive groups, not individual voters. The following proposition shows that $\mathcal{R}_{CC}$ has a low proportionality degree, while it is diverse.

Proposition 1.5. $\mathcal{R}_{CC}$ has a proportionality degree of ≤ 1 .

Proof. Fix k , $m = 2k$ and $n = \alpha k$ with $\alpha \in \mathbb{N}$. Divide the voters into k equal-size groups so that the voters from group i approve c_i and $\{c_{k+1}, \dots, c_{2k}\}$. Consider the following profile.

$$\begin{aligned} &|V_1| \times \{c_1\} \cup \{c_{k+1}, \dots, c_{2k}\} \\ &|V_2| \times \{c_2\} \cup \{c_{k+1}, \dots, c_{2k}\} \\ &\vdots \\ &|V_k| \times \{c_k\} \cup \{c_{k+1}, \dots, c_{2k}\} \end{aligned}$$

CC outputs all subsets from $\binom{C}{k}$ as winner sets. Note that for each $i \in [k]$, V_i is a ℓ -cohesive group as $\frac{n}{k} = \frac{\alpha k}{k} = \alpha$ and $V_i = \alpha$. When $W = \{c_{k+1}, \dots, c_{2k}\}$, each voters gets represented by k candidates. Hence, for each $j \in V_i$, $|A(j) \cap W| = k$. This means that:

$$\frac{1}{|V_i|} \cdot \sum_{j \in V_i} |A(j) \cap W| = \frac{1}{\alpha} \cdot \alpha \cdot k = k$$

Or, each voter gets represented by k candidates in each ℓ -cohesive group. When, $W = \{c_1, \dots, c_k\}$, however, each member of each ℓ -cohesive group gets represented by exactly 1 candidate. So,

$$\frac{1}{|V_i|} \cdot \sum_{j \in V_i} |A(j) \cap W| = \frac{1}{\alpha} \cdot \alpha \cdot 1 = 1$$

In other words, for each ℓ -cohesive group each voter gets represented by 1 candidate. Note that the definition of the proportionality degree states that the $f(\ell)$ must be a lower-bound. Hence, in this case, the number of representatives each voter from each ℓ -cohesive group get is 1. $\square$

Another way of showing the result from above, is by proving that $\mathcal{R}_{CC}$ satisfies justified representation. This can be found in Aziz et al. (2017).

This chapter is ended by making some remarks on the computational aspects of proportionality. Aziz et al. (2017) show that, given a committee W and an election $E = (A, k)$, it is coNP-complete to decide whether W satisfies EJR. Checking if a committee satisfies JR, however, is computationally easy. This can be done by checking for each candidate $c \in C$ if they are approved by a 1-cohesive group. Then, we check whether less than $\frac{n}{k}$ members of such a group are left without a representative. On the other hand, computing *some* committee that satisfies EJR can be done in polynomial time. This is shown by Aziz et al. (2017), but it can also be seen by noting that the MES and seq-Phragmén are polynomial-time computable and always return a committee satisfying EJR. Elkind et al. (2024) show that, computing a JR committee and maximising social welfare is NP-hard. This shows that combining an efficient and basic proportionality property with a welfare constraint is intractable.

1.4 Apportionment methods

An *apportionment problem* occurs typically when voters vote on parties, instead of individual candidates, and we have to fill a parliament *proportional* to the amount of votes each party has received. There are three big apportionment methods, which are used in almost all parliamentary elections over the world.

The context is given as follows: there are s political parties $P_1, \dots, P_s$. By n_i we refer to the number of voters of each party and k denotes the size of the winning committee (or parliament). Note that this is the same as the ABC context, but C consists of parties $\{P_1, \dots, P_s\}$ that contain candidates, instead of candidates only.

We start with the *D'Hondt* or *Jefferson method* named after Victor D'Hondt, a Belgian professor in law, and Thomas Jefferson, the American president (Pukelsheim, 2017).

Apportionment rule 1 (D'Hondt method). The number of rounds is equal to k . Let $x_i(r)$ denote the number of seats assigned to party P_i in round r . In each round r , calculate the quotient $\frac{n_i}{x_i(r)+1}$ and assign the r 'th seat to the party P_i with the highest quotient.

This apportionment method is part of the broader class of *divisor methods*. The following well-known apportionment rule also belongs to this class:

Apportionment rule 2 (Saint-Laguë method). The number of rounds is equal to k . Let $x_i(r)$ denote the number of seats assigned to party P_i in round r . In each round r , calculate the quotient $\frac{n_i}{2x_i(r)+1}$ and assign the r 'th seat to the party P_i with the highest quotient.

The following apportionment rule is not a divisor method.

Apportionment rule 3 (Largest remainder method). First, assign $\lfloor k \cdot \frac{n_i}{n} \rfloor$ to each party. Note that now $k - p + 1$ seats have been assigned. After this, assign the remaining $r < p$ seats to the parties with the largest remainders $k \cdot \frac{n_i}{n} - \lfloor k \cdot \frac{n_i}{n} \rfloor$ assigning each party at most one seat.

We have now framed the apportionment problem in terms of dividing seats to political parties. Note that it is also possible to replace political parties by ethnic groups, religious affiliations or any other social group consisting of candidates that deserve seats proportionally.

Some basic properties we want any reasonable apportionment rule to satisfy are given in the following definition. We write $F((n_1, \dots, n_s), k) = (x_1, \dots, x_s)$ to denote the output of method A applied to the s parties for k seats. Here x_i is the number of seats for P_i under F .

Definition 1.4 (Pukelsheim, 2017). Let F be an apportionment rule.

Anonymity $F((n_1, \dots, n_s), k) = F((n_{\pi(1)}, \dots, n_{\pi(s)}), k)$ for any permutation $\pi : [s] \rightarrow [s]$ on party names.

Balanced F is *balanced* if $n_i = n_j$ implies $|x_i - x_j| \leq 1$.

Concordance F is *concordant* if whenever $n_i > n_j$ we have $x_i \geq x_j$.

Decency F is *decent* if for all $c > 0$ it is the case that $A((n_1, \dots, n_s), k) = F((\frac{1}{c}n_1, \dots, \frac{1}{c}n_s), k)$

Weak proportionality F is *weakly proportional* (or *exact*) if whenever the quota of each party is an integer number, then the output of F is the vector of these integers.

Important to note is that *any* divisor method satisfies these five criteria. A proof of this fact can be found in Section 4.5 of Pukelsheim (2017). This means that, in particular, D'Hondt and Saint-Laguë also satisfy these five criteria. It also holds for the LRM method, even though this is not a divisor method.

The LRM method does not satisfy two other basic axioms. Instead of defining them formally, I explain the idea and refer the reader to Chapter 9 of the book by Pukelsheim (2017). It does not satisfy *house size monotonicity* or *voter monotonicity*. House size monotonicity captures the idea that an increase of k to $k + 1$ should not lead to a party losing seats. Voter monotonicity captures the same principle, but for increasing n to $n + 1$. D'Hondt and Saint-Laguë do satisfy these monotonicity axioms. In fact, any divisor method satisfies house-size monotonicity.

The last part of this section is devoted to the presentation of three divisor methods that each guarantee at least one seat for each party with at least one vote. These, together with D'Hondt's and the Saint-Laguë method are called the *five traditional divisor methods* by Balinski and Young (2001). Before the rules are mentioned, it must be noted that the output of any divisor method can be computed in linear time (Reitzig and Wild, 2024). This adds greatly to the appeal of this class of methods.

Apportionment rule 4 (Adams' method). In round 0, assign each party 1 seat, to prevent division by zero. The number of rounds is equal to $k - s$. Let $x_i(r)$ denote the number of seats assigned to party P_i in round r . In each round r , calculate the quorum $\frac{n_i}{x_i(r)}$ and assign the r 'th seat to the party P_i with the highest quorum.

This method favours smaller parties:

Apportionment rule 5 (Dean's method). In round 0, assign each party 1 seat. The number of rounds is equal to $k - s$. Let $x_i(r)$ denote the number of seats assigned to party P_i in round r . In each round r , calculate the quorum $\frac{(x_i(r) + \frac{1}{2}) \cdot n_i}{x_i(r)(x_i(r) + 1)}$ and assign the r 'th seat to the party P_i with the highest quorum.

This is an instance of a *divisor method with geometric rounding*.

Apportionment rule 6 (Huntington-Hill’s method). In round 0, assign each party 1 seat. The number of rounds is equal to $k - s$. Let $x_i(r)$ denote the number of seats assigned to party P_i in round r . In each round r , calculate the quotient $\frac{n_i}{\sqrt{x_i(r)(x_i(r)+1)}}$ and assign the r ’th seat to the party P_i with the highest quotient.

These rules assign each party one seat before the first round starts, to prevent dividing by 0. Table 1.1 contains the results of these apportionment method in the case there are three parties, with 85, 10 and 5 votes respectively. Suppose that $k = 5$. The results are given as a vector, where the i ’th entry corresponds to the assigned number of seats of the i ’th party. The parties are numbered according to their size, from largest to smallest.

Method:	D’Hondt	Saint-Laguë	LRM	Adams	Dean	Huntington-Hill
Output:	(4, 1, 0)	(4, 1, 0)	(4, 1, 0)	(3, 1, 1)	(3, 1, 1)	(3, 1, 1)

Table 1.1: Table of results for the various apportionment methods.

Chapter 2

Separate electorates

The notions of proportionality presented in the previous chapter are based on the approval ballots of the voters. These notions enforce that the final winner set reflects the *preferences* of the voters. In layman’s terms: they are based on *who you like* and not on *who you are*. The latter, representing voters on the basis of some external ‘label’ or ‘attribute’ — usually some religious, ethnic or socio-economic group they belong to — is a form of representation that has not been studied extensively in the ABC context (Lackner and Skowron, 2023). However, it is a form of representation that is (and has been) used in more than 35 countries (Tan and Preece, 2022). One example is Lebanon, where the lower house has designated seats for candidates from certain communities and the highest ranking positions, such as that of the president, are also reserved on the basis of religious and ethnic communities (Bogaards, 2019; Lijphart, 1969). In India, there are *reserved seats* in the national and regional legislatures for certain marginalised groups (Rai, 2005).

Although the literature on this form of representation in the ABC context is scarce, there is an approval based framework in which voters vote on attributes of the candidate. If the attributes are taken to be demographic traits, it is possible to reason about descriptive representation. The framework was developed by Kagita et al. (2021) and is called *attribute approval voting*.

In this section, I first argue why descriptive representation should be pursued. Then, a philosophical motivation for why we should consider external attributes of voters and not only of candidates will be presented. In fact, I will argue that ‘voter-centered’ representation provides a better balance between representing interests of voters and their demographic traits than ‘candidate-centered’ representation. As Kagita et al. (2021) do not add traits to voters in their framework, their framework will be criticised for the purpose of descriptive representation.

2.1 Descriptive and substantive representation

To motivate the representation of external attributes of voters, I present two important perspectives on political representation. The main sources I base this on are the papers by Ruedin (2020) and Pettit (2012). The two perspectives are called *descriptive* and *substantive* representation. The former emphasises the demographic similarity of the *representer* to the *representee*. So, the most important question for a proponent of this perspective is: does the representer have similar demographic properties as the person being represented? The latter emphasises that the representer should act in accordance with the *interests* of the representee. Immediately, there arises a strong argument in favour of substantive representation, which simultaneously critiques the descriptive variant. It is better to have

someone from a different group pass legislation that benefits your community than see someone from ‘your group’ join politics that does not represent your or your group’s interest. A tangential critique is that what is best for you as a member of the group might not be best for the group as a whole. Here, your (substantive) interests outweigh the interests you have as a member of a group.

A problem for this critique is that there is enough evidence to believe that regional and ethnic minorities are underrepresented in formal politics in almost all parts of the world (Ruedin, 2020). Descriptive representation might not always have the intended effect, but it is better than ignoring group membership. Some arguments based on empirical evidence can be found in the collection of Bird et al. (2010). The main takeaway being that representatives from a certain group are more likely to act in line with their group’s interests. One example is from Europe. There is evidence for the claim that ethnic minorities are more likely to ask immigration-related questions. Bird (2010) analyses voting behaviour of Canadian members of parliaments and finds that women are more likely to address women’s issues than men and ethnic minorities ask more questions relating to issues from their communities.

Furthermore, Casellas and Leal (2010) gather data from the United States and find that *[members of the House of representatives] were less likely, ceteris paribus, to exhibit conservative voting patterns when their districts had relatively large shares of Latinos*. It is important to note that conservatives are less likely to show voting behaviour that is in line with interests of minority groups. They find the same pattern whenever there are more African-American representatives. This pattern does not extend to women, however, and it is less significant when considering members of the Senate. Ruedin (2020) mentions that it indeed is the case that descriptive representation does not *necessarily* lead to substantive representation. But, he points out that there is a sizeable amount of evidence pointing towards a correlation between descriptive and substantive representation. The relation is that when there is no descriptive representation, the likelihood of substantive representation decreases significantly. Descriptive representation is not a necessary condition for substantive representation, but there is empirical data to regard it as a sufficient condition.

Another indication of sufficiency is that in situations of mistrust between representers and representees, descriptive representation can increase trust in the political system. She analyses trust in the context of African-American and women voters in the United States and finds that *representatives and voters who share membership in a subordinate group can also forge bonds of trust based specifically on the shared experience of subordination* (Mansbridge, 1999). For certain groups, this leads to representees contacting their representers and gaining trust in the political system. Hence, it may seem that descriptive representation, a superficial form of representation at face value, can lead to more substantial kinds of representation. On the other hand, Mansbridge (1999) also finds that different disadvantaged groups react differently to descriptive representation. Women, for instance, are not more likely to contact women representers.

In principle, whenever an individual enjoys substantive representation she enjoys a more ‘substantial’ form of representation than when she *only* enjoys descriptive representation; having someone work for your interests is more meaningful than when someone with the same background and experiences is in a position of power, but does not work in your interest. This stance seems to be rather superficial. As we have seen, there is a considerable amount of literature pointing towards a connection *from* descriptive representation *to* more substantial forms. It is important to note, however, that the arguments presented thus far are not irrefutable, but they are convincing enough to state that considering descriptive forms

of representation in ABC voting, where understanding proportionality is an important research goal, is not irrelevant.

The ABC model, however, is not rich enough to express group membership as each voter is only associated with a natural number. The forms of representation that have been formulated in this model, such as EJR, represent ‘preference-based-communities’, members of such groups do not have to share demographic traits. In fact, the model is not even able to express demographic traits. Preference-based representation bares more similarity to substantive representation. An example where this kind of representation is unsatisfactory is in the case of representing women. The preferences of women might not be cohesive in a lot of cases, but electing no women because there already exist candidates with similar properties does not seem to be right. In other words, in cases where men and women have the same preferences, we should not listen only to one of these groups, because their preferences are the same.

As mentioned, there exists a richer framework, developed by Kagita et al. (2021). Here, voters do not vote on candidates, but on *attributes* a candidate could have. This model can capture descriptive representation, if we take the attributes to be demographic traits and we assume that each voter votes for candidates with the same traits as themselves. This way, we do not have to add traits to the voters, since we can recover them based on their voting behaviour. We can now define the framework called *attribute approval voting* or *AAV*. Instead of $\mathcal{R}$, I will write $\mathcal{R}^{AAV}$ for AAV-voting rules.

Definition 2.1 (Attribute approval voting). Let N be a set of voters, and $C = \{c_1, c_2, \dots, c_m\}$ a set of candidates. Each candidate is associated with a d -dimensional attribute vector of d attributes. The attribute value of $c_i[j]$ of candidate c_i on dimension j is from a domain D^j such that $j \in [1, d]$. Let $A^j(i)$ denote the set of values in D^j approved by voter i . The goal is to select a committee W of size k candidates. We use $N(c_i[j])$ to denote the set of voters who have approved attribute-value $c_i[j]$.

For the remainder of this section, we only need to understand one AAV-voting rule, which is the AAV extension of $\mathcal{R}_{AV}$. Here we pick the candidate with the highest cumulative score over all properties, i.e., we select W by maximising $\sum_{i \in N} \sum_{j=1}^d |W^j \cap A^j(i)|$ where W^j is the winner with respect to j and $A^j(i)$ are voter i ’s approved attribute-values with respect to attribute j . The following example shows the working of the AAV model:

Example 2.1. A hospital wants to hire $k = 2$ people. Each person is associated with $d = 3$ attributes: prior experience in years (Exp), gender (G) and prior education (Edu). The domains are given as follows:

$$D^{\text{Exp}} = \{0, (0, 5), [5, \infty)\}, \quad D^{\text{G}} = \{M, F, X\} \quad \text{and} \quad D^{\text{PE}} = \{\text{Science, Arts, Engineering, Medical}\}$$

These are the candidates:

Candidate	Exp	G	Edu
a	0	F	Science
b	$[0, 5]$	F	Medical
c	0	M	Arts
d	$[5, \infty)$	X	Medical
e	0	M	Medical

To clarify the notion: $a[0] = a[\text{Exp}] = 0$ means that candidate a has attribute-value 0 for attribute 0, which in this case is ‘Exp’. This is the profile:

$$\begin{aligned}
&1 \times (\{[0, 5], [5, \infty)\}, F, \text{Medical}) \\
&1 \times (\{[0, 5]\}, F, \text{Engineering}) \\
&1 \times (\{0, [5, \infty)\}, \{M, F, X\}, \text{Medical})
\end{aligned}$$

We use AAV as our voting rule. There is a tie between $(0, 5)$ and $[5, \infty)$, F has the most votes and Medical also has the most votes. This means that $\{b, d\}$ is the winner set. $\diamond$

The advantage of this definition is that it requires a relatively small modification of the regular system of approval voting, which makes it easier to understand. The downside is that it leads to some counter-intuitive results since we do not vote on candidates *and* attributes, but only on attributes. The problem with this is that there might not exist a candidate that has all the properties you would like. It is then hard to decide what a good second option is. Consider the following example that supports this conceptual critique.

Example 2.2. Suppose that we are voting on three attributes for a presidential election. The attributes are ‘religion’, ‘constituency’ and ‘gender’. You vote for $(\text{Muslim}, \text{Rotterdam}, \text{woman})$. But the candidates are: $\{(\text{Muslim}, \text{Rotterdam}, \text{man}), (\text{Christian}, \text{Rotterdam}, \text{woman}), (\text{Muslim}, \text{Doetinchem}, \text{woman})\}$. Doetinchem is a rural city in the Netherlands. It is not obvious that these candidates have an equal appeal to you. They all satisfy two of your attributes, but a Christian woman from Rotterdam seems to be further removed from you than a Muslim man from Rotterdam. $\diamond$

For one voter, it could be the case that their Muslim identity is more important than the constituency of the voter. For another voter, the gender of a candidate is the most important factor. A solution for this problem could be a ‘weighted’ AAV model, in which each voter assigns a weight to each attribute. This kind of voting, however, requires a lot of information from the voters. They need to vote on multiple attributes and also compare each attribute. This conflicts with the appeal of the simplicity of ABC voting.

There is another conceptual critique of this model. This model gives counter-intuitive results in the case there is a group with one demographic characteristic but with diverse preferences. Before I present this critique, I first introduce the notion of *simple justified representation* (SJR) from Kagita et al. (2021) work. This is one of the two notions of justified representation they present, the other being *compound justified representation* (CJR). The latter is a stronger notion, since CJR implies SJR. The idea behind SJR is that if a committee is large enough ($\geq \frac{n}{k}$) and they are cohesive with respect to one attribute, then they deserve representation with respect to (any) one attribute. I first

define *attribute cohesiveness* analogous to how we defined ℓ -cohesive group. The first condition is a size constraint and the second condition captures the idea that the group has the same preference with respect to at least one attribute.

Definition 2.2. A group $V \subseteq N$ is *attribute-cohesive* (with respect to attribute j) if

1. $|V| \geq \frac{n}{k}$.
2. There exists a $j \in [1, d]$ such that $\bigcap_{i \in V} A^j(i) \neq \emptyset$.

We can now define SJR:

Definition 2.3. An AAV rule $\mathcal{R}^{AAV}$ satisfies simple justified representation, if for each election instance $E = (A, k)$, each winning committee $W \in \mathcal{R}^{AAV}(E)$ and each attribute cohesive group of voters V there exists a voter $i \in V$ such that they are represented with respect to an attribute $j' \in [1, d]$, i.e., $\exists j' \in [1, d]$ such that $\bigcup_{i \in V} A^{j'}(i) \cap W^{j'} \neq \emptyset$.

Informally, the difference between SJR and CJR is that CJR wants voters to get represented with respect to the attribute they share. We now consider an example in which SJR is satisfied, but the winning set does not seem representative on an intuitive level.

Example 2.3. Suppose there are six voters approving of three attributes (*gender, ethnicity, income*). Let $k = 2$. For simplicity, assume that the first three voters are men (those who approve of *man*) and the other half consist of women. ‘Hindostani’s’ or ‘Indo-Surinamers’ are a group of people that mainly live in the Netherlands and Suriname who trace their ancestry to the Indian subcontinent. They are one of the larger minorities in the Netherlands, with significant communities in the Randstad (Oudhof et al., 2011). This is the profile after voting:

$$\begin{aligned} & 3 \times (\text{man}, \text{Hindostani}, \text{high}) \\ & 2 \times (\text{woman}, \text{Hindostani}, \text{low}) \\ & 1 \times (\{\text{woman}, \text{man}\}, \text{Dutch}, \{\text{low}, \text{high}\}) \end{aligned}$$

The candidates are: $\{(\text{man}, \text{Hindostani}, \text{low}), (\text{woman}, \text{Dutch}, \text{high}), (\text{man}, \text{Hindostani}, \text{high})\}$. Both the women and the men are attributes-cohesive with respect to gender. A winning committee containing $(\text{man}, \text{Hindostani}, \text{low})$ and $(\text{man}, \text{Hindostani}, \text{high})$ satisfies SJR. In the case of women, let j' be *income*. We see that $\bigcup_{i \in \text{women}} A^{\text{income}}(i) \cap W^{\text{income}} = \bigcup_{i \in \text{women}} \{\text{low}\} \cap \{\text{low}\} \neq \emptyset$. So, the group of women gets their deserved representation in the form of a low-income candidate. The men get a candidate that satisfies all their desired attributes, so it satisfies SJR immediately. CJR would lead to no outcome, since the group of women is cohesive with respect to both income and gender, but there exists no candidate that is both low-income and a woman. $\diamond$

Whereas it is possible to argue for EJR from the perspective of substantive representation, but not from the perspective of descriptive representation, neither perspective can be used to argue for SJR. From the example we see that women can be represented by a low-income candidate. This obviously does not satisfy the descriptive variant and it also does not satisfy the substantive variant, since it is not necessarily the case that the interests of low-income people align with those of women. CJR, however, is a more appealing form of representation. If we assume that voters approve of traits that they share, then CJR satisfies the descriptive variant. The problem with this is that Kagita et al.

(2021) show that the AAV variant of PAV does not satisfy SJR for $k \geq 3$ and $d \geq 2$ or $d \geq 3$ and $k \geq 2$. This also means that it does not satisfy CJR for very small parameters.

This leads to a twofold critique of the AAV model with respect to descriptive representation: (i) there is an interplay between different attributes that is not captured by the model as presented in Example 2.2, (ii) SJR, one simple proportionality axiom, does not lead to descriptive (or substantive) representation and (iii) CJR, the most natural proportionality axiom, is provably unsatisfactory.

This critique can be extended by the fact that AAV is ‘candidate-centered’, whereas descriptive representation has a ‘voter-centered’ reading. I present this critique by making use of an example: *A country split into two communities A and B is voting for a president and a prime minister. There are three candidates, a_1 and a_2 from A and b from B.* The simplest way to guarantee descriptive representation is by fixing that the president is from A and the prime minister is b . But, this can easily lead to the situation critiques of descriptive representation point towards: it could be that b does not represent the interests of B, but a_1 does, and in that situation all seats should be filled by candidates from A.

One way to circumvent this is by having a ‘voter-centered’ system, in which the voter is represented on the basis of her community, but she decides who to vote for. In the case of our toy-example, we could fix that people from A vote for the president and people from B for the prime minister. Desiring proportionality on the candidate level, which can be captured in the AAV model, can lead to situations in which the interests of the group you are trying to represent get ignored. Instead, an approach in which groups get the *power* to decide a part of the final outcome balances substantive and descriptive forms of representation. If we do not do this and we force the final winner set to consist of candidates from a certain community, it could be the case that the community does not approve of her.

2.2 Two perspectives on separate electorates

Until now, the arguments for a voter-centered model have been of a conceptual and philosophical nature. Another advantage of a model in which voters get represented is that the formalisation requires only a small modification of the existing ABC model. AAV, on the other hand, requires a more hefty redefinition. Before the new model is presented formally, the intended working is showcased by using the running example.

Example 2.4. Recall that $n = 100$, $m = 6$ and $k = 2$. Since we have not formalised *group* or *community membership*, the horizontal line indicates the community structure. In words, the 51 voters approving $\{l, b\}$ belong to one community and the remaining 49 voters constitute another community.

$51 \times \{l, b\}$
$18 \times \{c, s\}$
$16 \times \{g, s\}$
$15 \times \{p, s\}$

One straightforward way to enforce representation of both communities is by assigning them a fixed number of the final winner set W . This can be done by running two separate elections (N_1, q_1) and (N_2, q_2) . Here N_i refers to the groups and q_i refers to the number of seats that group is assigned. I refer to q_i as the *weights* per community. Since both groups are roughly equal, it makes to sense to fix $q_1 = q_2 = 1$. For simplicity, we use AV as our rule for both elections. After running two

separate elections per community, we obtain $W_1 = \{\{l\}, \{b\}\}$ and $W_2 = \{s\}$. We still have to consider a systematic way to combine community winner sets in the general setting, but in this example it makes sense to combine them to form $W = \{\{l, s\}, \{b, s\}\}$. $\diamond$

Allotting the seats can be seen as apportionment for communities, instead of parties. A downside of this approach is the necessity to round the number of seats. In the running example, it is easy to justify why both communities deserve half of the seats. For instance, in the case where there is one community consisting of 90 and one consisting of 10 voters, it seems unfair to allot one seat to each. To circumvent this problem, instead of reserving seats based on community-size *relative* to something else — the size of the winner set, in the example above —, we can add weights to each community based on their *absolute* size. As an example, consider:

Example 2.5. Recall:

$$\begin{array}{r} 51 \times \{l, b\} \\ \hline 18 \times \{c, s\} \\ 16 \times \{g, s\} \\ 15 \times \{p, s\} \end{array}$$

In each community, run separate elections using $\mathcal{R}_{AV}$ for the full k . $\{l, b\}$ wins in the community above and $\{c, s\}$ in the community below. Now, apply $\mathcal{R}_{PAV}$ to the following profile:

$$\begin{array}{r} 51 \times \{l, b\} \\ 49 \times \{c, s\} \end{array}$$

Then, return the winners of this election, that is, $\{\{l, s\}, \{l, c\}, \{b, c\}, \{b, s\}\}$.

Note that the election is divided into two separate steps. In the first step we acquire information about the most popular candidates in each community. Then, the final winner set is determined by applying a voting rule to the winners of the communities. If the last rule is proportional, the final outcome will also be proportional. $\diamond$

The next section is devoted to formalising descriptive representation in ABC voting and two classes of voting rules.

2.3 The model

To define the exemplified social choice setting, we briefly recall the setting of ABC voting. Recall that N refers to the set of voters, C to the candidates, k to the final winner set, A to a profile and $E = (A, k)$ is an election. We define a *community structure* on N as a partition of N in to s disjoint sets. From now on, $\mathcal{Q}$ refers to a partition of N and s refers to $|\mathcal{Q}|$, i.e., how many communities there are. So, $\mathcal{Q} = \{N_1, N_2, \dots, N_s\}$ such that $\bigsqcup \mathcal{Q} = N$ and $N_i \cap N_j = \emptyset$ if $i \neq j$. I will also write $\mathcal{Q} = [|N_1|, |N_2|, \dots, |N_s|]$ to refer to communities. We assume that $s \leq k$. A community ballot for a community $Q \in \mathcal{Q}$ is defined as a restriction on the function A to $A|_Q: Q \rightarrow \wp(C)$. Lastly, I name this setting *ABC voting with communities*.

Note that the assumption that $s \leq k$ need not be necessary. If $s > k$ it might be the case that some communities get no seats. Thus, no representation. To prevent this from happening, we assume that $s \leq k$. The interpretation of the community structure is informal; the partition can be based on

any demographic or social trait. It is possible to formalise this analogously to how AAV is formalised. Instead of assigning attributes to candidates, we would assign candidates to voters and partition N based on these attributes. Consider the following example in this setting:

Example 2.6. Recall Example 1. We extend this example and add more allergic friends. In total, a group of $n = 10$ friends is having lunch. They are partitioned as follows: $\mathcal{Q} = [3, 2, 2, 3]$. Recall that the i 'th entry corresponds to the size of N_i . N_1 consists of friends allergic to fava beans, N_2 consists of friends with a gluten allergy and N_3 is a group with an allergy to apples. N_4 consists of the non allergic group. $\diamond$

I define a *sub-election* or a *community-election* as an election instance (N_i, q_i) , where q_i are the *weights* per community. $\mathbf{q} = (q_1, \dots, q_s)$ refers to the vector of weights per community. An *election instance with communities* is written as $E = (A, \mathcal{Q}, k)$ where k is the desired committee size, A the approval ballot and $\mathcal{Q}$ the community structure. I will not use ‘voting rule’ to refer to rules in this social choice setting, but ‘voting procedures’ to highlight the difference. Furthermore, I will not use $\mathcal{R}$ but $\mathcal{R}^\rho$ for ABCC rules.

We can now formally define the procedure from Example 2.5. Recall that an ABCC procedure takes as input an election $(A, \mathcal{Q}, k)$ and returns winner sets from $\binom{C}{k}$. First, there is an informal formulation and then a formulation as a TSP. We call it the *naively weighted procedure (NWP)*.

Definition 2.4 (Naively weighted procedure (NWP)). The input of this procedure, $\mathcal{R}_{NWP}^\rho$, is an ABCC election $(A, \mathcal{Q}, k)$, a tuple of s voting rules $(\mathcal{R}_1, \dots, \mathcal{R}_s)$ and a special voting rule $\mathcal{R}_{s+1}$. The first step is to find the weights, for NWP these are given by: $\mathbf{q} = (|N_1|, |N_2|, \dots, |N_s|)$. That is, for each community N_i for $i \in [s]$ their weight is their size. Now, run $\mathcal{R}_i(A_i, k)$ for each $i \in [s]$ to obtain the community winner sets. The last step consists of aggregating these winner sets into final winning committees. For this, we have to construct new profiles and apply $\mathcal{R}_{s+1}$ to each of these profiles. The profiles are defined by taking each possible combination of profiles, i.e., an element $x \in \times_{i \in [s]} \mathcal{R}_i(A_i, k)$ and constructing $|N_i|$ ballots that contain the candidates from $x[i]$, where $x[i]$ denotes the i 'th entry of x corresponding to a winner set from $\mathcal{R}_i(A_i, k)$, and adding these to the profile Π . All such profiles to the set of profiles $\mathbf{A}$. Now, apply $\mathcal{R}_{s+1}$ to (Π, k) for all $\Pi \in \mathbf{A}$, and return all the set of all winner sets of this rule.

We now define the procedure from Example 2.4. It will be named the *reserved seats procedure (RSP)*. Before defining it, we consider a looming problem. Combining the results of two disjoint elections is not always as straightforward as in the previous election. This becomes clear in this example:

Example 2.7. Let $n = 8$, $C = \{a, b, c, d\}$, $k = 3$, $\mathcal{Q} = \{N_1, N_2\}$ and $\mathbf{q} = (2, 1)$ such that this is the outcome:

$$\begin{array}{r} 3 : \{a, b\} \\ 2 : \{a, c\} \\ \hline 2 : \{a, d\} \\ 1 : \{a\} \end{array}$$

If we again run AV internally we get winner set $\{a, b\}$ and $\{a\}$, but how do we combine these into a winner set of size k ? This is the question we investigate now. $\diamond$

Both RSP and NWP consist of three steps. First, determine the weights $\mathbf{q}$. Second, find the winners of the community elections. Lastly, use the winners of these elections to construct the final winner sets. Therefore, we will refer to these procedures as *two-step procedures*. For RSP, determining how many weights each community deserves can be seen as an apportionment problem. Instead of s political parties with n_i voters, there are s communities with n_i members. The number of seats is k and in each round a community gets assigned a seat. We want to ensure that each community gets at least one seat, so we must use an apportionment method that does this. Recall that Huntington-Hill, Adams and Dean guarantee this. Generally speaking, we call such an apportionment method *diverse*. Recall that we assumed that $s \leq k$.

We can now define $\mathcal{R}_{RSP}^\rho$:

Definition 2.5. RSP takes as input an ABCC election $(A, \mathcal{Q}, k)$ and a tuple of voting rules ρ . First, apply a diverse apportionment method to $((n_1, \dots, n_s), k)$ where $n_i = |N_i|$, for all $i \in [s]$, to obtain the weights $\mathbf{q}$. Then, run $\mathcal{R}_i(A_i, q_i)$ where q_i is the i 'th entry of $\mathbf{q}$ for each $i \in [s]$. Construct $\times_{i \in [s]} \mathcal{R}_i(A_i, q_i)$. We now want to return every possible combination of community winner sets as final winner sets. So, for each $x \in \times_{i \in [s]} \mathcal{R}_i(A_i, q_i)$ take its union $\bigcup x$ and check if there exists at least one x such that $|\bigcup x| = k$. If this is the case, return only the $\bigcup x$ that are of size k . Otherwise, if all $\bigcup x$ contain fewer elements than k , extend each of the sets $\bigcup x$ with every possible set — which are called ‘extensions’ — from $\binom{C \setminus \bigcup x}{|\bigcup x| - k}$, i.e., extend each $\bigcup x$ with all possible combinations of remaining candidates such that their union is of size k .

Both of these procedures can be formalised as algorithms. For this, the reader is referred to the appendix; Algorithm 1 corresponds to NWP and Algorithm 2 to RSP. We apply both procedures to the tricky example above (2.7), to show their working.

Example 2.8. Let $n = 8$, $m = 4$, $k = 3$ and $\mathcal{Q} = [5, 3]$. Consider:

$$\begin{array}{l} 3 : \{a, b\} \\ 2 : \{a, c\} \\ \hline 2 : \{a, d\} \\ 1 : \{a\} \end{array}$$

We use $\mathcal{R}_{AV}$'s only for all voting rules R_i , $i \in [s + 1]$.

First, we show $\mathcal{R}_{RSP}^\rho$. For the apportionment method, we take Huntington-Hill and obtain $\mathbf{q} = (2, 1)$. $\{a, b\}$ is the winner set of the first community election and $\{a\}$ for the second community. Now, construct $\{a, b\} \times \{a\} = \{(a, b, a)\}$. Taking the union leads to $\{a, b\}$. For the last step, we extend this set to $\{a, b, c\}$ and $\{a, b, d\}$ and return these as the winner sets.

Now, we show $\mathcal{R}_{NWP}^\rho$. Note that $\{a, b, c\}$ wins in the first community and there is a tie between $\{a, b, d\}$ and $\{a, c, d\}$ in the second community. We construct two profiles:

$$\begin{array}{l} 3 \times \{a, b, c\} \\ 2 \times \{a, b, d\} \end{array}$$

$$\begin{array}{l} 3 \times \{a, b, c\} \\ 2 \times \{a, c, d\} \end{array}$$

Applying $\mathcal{R}_{AV}$ in both cases leads to $\{a, b, c\}$ winning. $\diamond$

2.4 Observing the two procedures

Until now, we have defined the two procedures and gained some intuition based on some simple examples. To further understand how these procedures work, I systematically generated all possible profiles for different combinations of $n \in \{1, 2, \dots, 8\}$, $m \in \{1, 2, \dots, 6\}$, $k \leq m$ and partitions $\mathcal{Q}$ of N . Next to this, both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$ were applied to each profile with ρ consisting of $\mathcal{R}_{AV}$'s only. The reasoning behind this is twofold: (i) it can be computed in polynomial time and (ii) since we already enforce proportionality through the structure of the model and we want to understand how well it does this, using a simple rule that is not proportional itself gives a good picture of for which community structures the procedure succeeds in being proportional and when it does not.

One voter can pick from 2^m sets of candidates and there are n voters. So, there are 2^{m^n} possible profiles. This means that the number of possible profiles grows doubly exponential in n and m . For $n = 3$ and $m = 6$, we would have to check 2^{18} different options, which is undoable. Furthermore, applying $\mathcal{R}_{NWP}^\rho$ leads to running s community elections for winner sets of size k . So, if the community winner set is smaller than k , the algorithm adds all other remaining candidates. To prevent duplicate elections, I restricted to *knapsack voting* (Goel et al., 2019). This means that each voter is allowed to submit ballots of size k only. This way, we ensure that the final winner set is of size k and we prevent $\mathcal{R}_{NWP}^\rho$ from duplicating outputs because of tie-breaks in the last stages. Remark that this leaves us with $\binom{m}{k}^n$ total profiles. For $n = 3$, $m = 6$ and $k = 3$ this means that we check $20^3 = 8000 \approx 2^{12.97}$ profiles, which is a significant improvement.

To get an overview of how well these two procedures function, we want to answer two questions: (i) does the procedure satisfy EJR? If not, what is the probability that there is a winner set W in the output that does not satisfy EJR? (ii) Does it not satisfy EJR when AV does? And (iii) how do the two procedures compare given questions (i) and (ii)? The general conclusion after generating all possible profiles and running the two procedures, is that RSP performs better with respect to both metrics. There is only one instance, when $n = 3$, $m = 6$ and $k = 4$, that RSP does not satisfy EJR, but NWP does. In this section, I give a *qualitative* overview of the two procedures, whereas Appendix A contains a *quantitative* overview.

First, I give an overview of the parameters $(n, m, k, \mathcal{Q})$ for which neither procedure satisfies EJR, NWP does and RSP does not satisfy EJR and vice versa.

RSP does not satisfy EJR, NWP does	$(3, 5, 3, [1, 1, 1]), (3, 6, 4, [1, 1, 1])$
NWP does not satisfy EJR, RSP does	$(3, 4, 2, [1, 2]), (3, 5, 2, [1, 2]), (3, 6, 2, [1, 2]),$ $(4, 4, 2, [1, 3]), (4, 5, 2, [1, 3]), (5, 4, 2, [1, 4])$
Neither procedure satisfies EJR	$(3, 5, 3, [1, 2]), (3, 6, 3, [1, 2]), (3, 6, 3, [1, 1, 1]),$ $(4, 4, 2, [2, 2]), (4, 5, 2, [2, 2]), (4, 5, 3, [2, 2]), (4, 5, 3, [1, 3])$

Per combination of parameters, we show why RSP or NWP does not satisfy EJR.

RSP does not satisfy EJR, NWP does

I first show some examples of when RSP does not satisfy EJR, but NWP does. This is the rarest case. Note that for the determination of the weights the Huntington-Hill method was used.

Example 2.9. Let $n = 3$, $m = 6$, $k = 4$ and $\mathcal{Q} = [1, 1, 1]$.

$1 \times \{a, b, c, d\}$
$1 \times \{a, b, c, e\}$
$1 \times \{a, b, c, f\}$

For $\mathcal{R}_{RSP}^\rho$, the weights are $\mathbf{q} = (2, 1, 1)$. Here, we broke ties in favour of the first community. Note that any winner set that contains $\{a, b, c\}$ would satisfy EJR. But we see that $\{b, d\} \cup \{e\} \cup \{f\}$ is one of the winners, so $\{b, d, e, f\}$ is one of the winner sets, which does not satisfy EJR. NWP just redoes a regular AV-election and has the three ballots as elements of the winner set. $\diamond$

The other case where RSP does not satisfy EJR, but NWP does is $(3, 5, 3, [1, 1, 1])$ and it is similar to this one. The difference is that each voter submits a ballot of size 3 instead of 4 and they all approve of $\{a, b, x\}$, where x is different for each community. Put differently, we obtain this profile by removing c from all approval ballots from the example above. It is clear that in that case $\{d, e, f\}$ wins, which again does not satisfy EJR.

Although there is a probability that one of the winner sets from $\mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$, for the $(A, \mathcal{Q}, k)$ of the previous example, might not satisfy EJR, it is still possible to argue that this outcome is proportional. Whenever $\{d, e, f\} \subset W$, we do not only have representatives for each voter but a form of *unique* representation, since each voter has one representative that is not shared by someone else. This form of representation is similar to *perfect representation*, which captures the idea that if the electorate can be divided into $k \frac{n}{k}$ -sized groups in such a way that each group gets assigned a candidate, then a rule should do this. It is known that this form of representation is incompatible with EJR (Sánchez-Fernández et al., 2017). We discuss this in Chapters 3 and 4.

NWP does not satisfy EJR, RSP does

We now discuss the case in which NWP does not satisfy EJR, but RSP does. We will see that RSP passes a more intuitive verdict in all these cases. Consider this first example:

Example 2.10. Let $n = 3$, $m \in \{4, 5, 6\}$, $k = 2$ and $\mathcal{Q} = [1, 2]$.

$1 \times \{a, b\}$
$1 \times \{a, b\}$
$1 \times \{c, d\}$

We first do NWP. The winner set in the first community is $\{a, b\}$ in N_1 and in N_2 all elements of $\binom{C}{2}$ win. Since $\{c, d\}$ is such a winner and since the weights of the second community always outweigh the results of the first community, this election boils down to domination of N_2 . There is a probability of $\frac{1}{6}$ that $\{c, d\}$ wins, which does not satisfy EJR. RSP outputs all Cartesian products of $\{\{a\}, \{b\}\}$ and $\{\{a\}, \{b\}, \{c\}, \{d\}\}$. Hence, $W = \{\{a, c\}, \{a, d\}, \{a, b\}, \{b, c\}, \{b, d\}\}$. These all satisfy EJR. $\diamond$

Here it becomes apparent that whenever one community is strictly larger than any other community and the last rule we apply is $\mathcal{R}_{AV}$ this leads to some form of ‘domination’ of this community. This is especially the case if the other communities are not cohesive.

I wrote $m \in \{4, 5, 6\}$ in the example above since all of the counterexamples are similar; there are two voters that share at least one candidate and there is one ‘disjoint voter’. This disjoint voter is in

the bigger community with one of the other voters. Because of how the voting rule works, there is a chance that the disjoint voter gets their whole approval ballot as the winning set. Note furthermore that $\mathcal{R}_{AV}$ applied to the example above would lead to $\{a, b\}$, which fairer (in terms of EJR) than $\mathcal{R}_{NWP}^\rho$. So, we see that this procedure might even lead to outputs that are *less* proportional than AV.

We now consider $(4, 4, 2, [1, 3])$. There are 17 profiles for which EJR is never satisfied. This is the general structure:

Example 2.11. Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = [1, 3]$. Consider this profile:

$$\begin{array}{c} 1 \times \{a, b\} \\ \hline 1 \times \{a, b\} \\ 2 \times \{c, d\} \end{array}$$

To satisfy EJR, any element from $\{a, b\} \times \{c, d\}$ would suffice. NWP, however, outputs $\{c, d\}$. This is again a case of ‘dominance’ because of the size of one of the communities. RSP, however, would output any element from $\{a, b\} \times \{c, d\}$, since the weights have to satisfy diversity. $\diamond$

The cases for $(4, 5, 2, [1, 3])$ and $(5, 4, 2, [1, 4])$ are similar. The general idea is that the community structure does not align with the structure of the preferences. Cohesion gets ignored, in favour of descriptive representation:

Example 2.12. Let $n = 4$, $m = 5$, $k = 2$ and $\mathcal{Q} = [1, 3]$. Consider this profile:

$$\begin{array}{c} 1 \times \{a, b\} \\ \hline 1 \times \{a, c\} \\ 1 \times \{b, d\} \\ 1 \times \{c, d\} \end{array}$$

$\{c, d\}$ is one of the winner sets here and it does not satisfy EJR, since there is one 1-cohesive group, $N(a)$, that gets no representation. $\diamond$

Neither procedure satisfies EJR

Now, I present some cases in which neither procedure satisfies EJR. For $(3, 5, 3, [1, 2])$ NWP still performs worse, percentage wise. The reasons why EJR is not always satisfied for $(3, 6, 3, [1, 2])$ are similar.

Example 2.13. Let $n = 3$, $m = 5$, $k = 3$ and $\mathcal{Q} = [1, 2]$.

$$\begin{array}{c} 1 \times \{a, b, c\} \\ \hline 1 \times \{a, b, d\} \\ 1 \times \{c, d, e\} \end{array}$$

NWP outputs $\{c, d, e\}$ with a probability of $\frac{1}{6}$. For RSP with $\mathbf{q} = (1, 2)$, the output is $\{\{a\}, \{b\}, \{c\}\} \times \{\{a, d\}, \{b, d\}, \{c, d\}, \{d, e\}\}$ which also includes $\{c, d, e\}$. $\diamond$

The reasons for not satisfying EJR are the same as before. Belonging to a community does not align with the voters’ preferences. For $(3, 6, 3, [1, 1, 1])$ the reasons for failing EJR are different for the two procedures. I only show RSP, as NWP returns the same as AV does for this example.

Example 2.14. Let $n = 3$, $m = 6$, $k = 3$ and $\mathcal{Q} = [1, 1, 1]$.

$$\begin{array}{r} 1 \times \{a, b, c\} \\ 1 \times \{a, b, d\} \\ \hline 1 \times \{a, c, e\} \end{array}$$

$\{a, b, c\}$ satisfies EJR and everyone is represented by two candidates. Note that this is the output of $\mathcal{R}_{AV}$. RSP, however, takes a big detour. First, it returns three singletons per community as the community winner sets and then it aggregates the outputs into final winner sets. The reason for this is that the ballots are larger than the seats ($q_i = 1$). One of the outputs is $\{c, d, e\}$, which does not satisfy EJR. $\diamond$

Again, it can be argued that $\{c, d, e\}$ is proportional. The following is a particularly interesting case:

Example 2.15. Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = [2, 2]$. Consider:

$$\begin{array}{r} 1 \times \{a, b\} \\ 1 \times \{c, d\} \\ \hline 1 \times \{a, b\} \\ 1 \times \{c, d\} \end{array}$$

Both RSP and NWV make detours, which lead to the same output as AV, NWP recognises all six $\binom{4}{2}$ 2-subsets of C as winners per community, and then runs all 36 combinations to come to the conclusion that all the 2-subsets are winners. RSP does essentially the same, but it outputs all the Cartesian products of size 2 of the singletons. In both cases, $\{a, b\}$ and $\{c, d\}$ are included in the final winner set. These do not satisfy EJR. $\diamond$

In the case $(4, 5, 3, [2, 2])$ NWP, again, leads to a counter-intuitive result:

Example 2.16. Let $n = 4$, $m = 5$, $k = 3$ and $\mathcal{Q} = [2, 2]$.

$$\begin{array}{r} 1 \times \{a, b, c\} \\ 1 \times \{a, b, c\} \\ \hline 1 \times \{a, b, c\} \\ 1 \times \{a, d, e\} \end{array}$$

$\{a, b, c\}$ satisfies EJR. NWP, however, also outputs $\{a, d, e\}$ since there is a case in which both $\{a, d, e\}$ and $\{a, b, c\}$ get 2 votes and by tie-breaking $\{a, d, e\}$ is returned. $\diamond$

But, RSP also leads to a counter-intuitive result for these parameters:

Example 2.17. Let $n = 4$, $m = 5$, $k = 3$ and $\mathcal{Q} = [2, 2]$.

$$\begin{array}{r} 1 \times \{a, b, c\} \\ 1 \times \{a, d, e\} \\ \hline 1 \times \{b, c, d\} \\ 1 \times \{b, c, d\} \end{array}$$

Because of tie-breaking, one of the communities gets one extra weight. If this goes to the first community, $\{a, d, e\}$ is one of the winner sets. $\diamond$

Lastly, note that the problematic cases for $(4, 5, 3, [1, 3])$ are similar to the problematic cases for $(4, 4, 2, [1, 3])$ and $(4, 5, 3, [1, 3])$. Therefore, I omit them.

More profiles

We finish by presenting two examples that were not part of the systematic generation, but they do highlight two classes of profiles for which the two procedures do not satisfy EJR. Recall that in Example 2.15 we saw that the number of elections that are done in the last step explode for NWP. Another example of this is the following:

Example 2.18. Let $n = 5$, $m = 5$, $k = 3$ and $\mathcal{Q} = [2, 3]$. We want to select a committee of size $k = 3$. The approval profile is as follows:

1 : {a, b, c}
2 : {a, b, d}
3 : {a, b, e}
4 : {d, e, f}
5 : {c, e, f}

There is a tie between $\{a, b, c\}$ and $\{a, b, d\}$ in N_1 so their ballots stay the same. In N_2 we have the following winner sets

$$\{\{c, e, f\}, \{d, e, f\}, \{a, e, f\}, \{b, e, f\}\}$$

Following this procedure, we have to run 2×4 elections.

$2 \times \{a, b, c\}$	$2 \times \{a, b, c\}$	$2 \times \{a, b, c\}$	$2 \times \{a, b, c\}$	$2 \times \{a, b, d\}$	$2 \times \{a, b, d\}$	$2 \times \{a, b, d\}$	$2 \times \{a, b, d\}$
$3 \times \{c, e, f\}$	$3 \times \{d, e, f\}$	$3 \times \{a, e, f\}$	$3 \times \{b, e, f\}$	$3 \times \{c, e, f\}$	$3 \times \{d, e, f\}$	$3 \times \{a, e, f\}$	$3 \times \{b, e, f\}$

Hence, $W = \{\{c, e, f\}, \{d, e, f\}, \{a, e, f\}, \{b, e, f\}, \}$.

RSP recognises $\{a, b\}$ as the winner set in N_1 and $\{e\}$ as the winner set in N_2 . So, $\{a, b, e\}$ is the final winner set. $\diamond$

Consider:

Example 2.19. Let $n = k = 4$ and $m = 5$.

1 : {a, b}
2 : {c}
3 : {d, e}
4 : {a, b}

For $V = \{1, 4\}$ we get $|V| = 2 \leq 2 \cdot \frac{4}{4}$ and $\bigcap_{i \in V} |A(i)| = 2$ so V is a 2-coalition. But, under (i), $\{c, d, e\} \cup \{a\}$ could be a winner. So, it does not always satisfy EJR. $\diamond$

All profiles for which RSP or NWP do not lead to a winner set that satisfies EJR have one thing in common: the community structure have less *internal cohesiveness* than *external cohesiveness*. I explain the phenomenon informally in this section, but the following chapters include a formal analysis of this observation. In each case a significant proportion of a community shares more candidates with voters from other communities. For instance, in Example 2.10 there is a 1-cohesive group that gets divided and ignored by the community structure. This leads to the following hypothesis: *whenever the internal cohesion is larger than the external cohesion, EJR is satisfied by RSP*. In Example 2.19 we see that an ℓ -coalition gets divided over two communities. There is a form of *underrepresentation* of a deserving community. There is also a form of *overrepresentation*:

Example 2.20. Let $n = 100$, $m = 10$, $k = 6$ and $\mathcal{Q} = [5, 10, 85]$. Note that $\mathbf{q} = (1, 1, 2)$.

$5 \times \{a\}$
$10 \times \{b\}$
$84 \times \{c, d, e, f\}$

$V = \{16, \dots, 99\}$ deserves three candidates, since $|V| = 84 > 3 \cdot \frac{10}{4}$. But the output is $\{a, b\} \cup X$, where $X \in \binom{\{c, d, e, f\}}{2}$. $\diamond$

Here, we see that an ℓ -coalition does not get assigned enough weights. There is overrepresentation of a community that deserves nothing and underrepresentation of a deserving coalition. Note that JR is satisfied.

The general pattern of these can be summarised as follows:

- If there is a form of under- or overrepresentation in terms of weights, RSP performs bad. We saw this in the previous two examples.
- In cases where the community structure does not align with the structure of the voters – voters share more candidates with voters outside of their community than inside – NWP and RSP both perform bad.
- There are cases in which RSP does not satisfy EJR, but it does seem proportional.

We investigate these observations from a more theoretical viewpoint in Chapter 4.

Chapter 3

Axiomatic and computational properties

This section contains a presentation of the most important axioms from the literature on ABC voting and a redefinition for the ABCC context. The guiding sources for this chapter are the overview of multi-winner axioms by Elkind et al. (2017) and the booklet by Lackner and Skowron (2023). This chapter is divided into five sections, each section containing a different class of axioms: (3.1) basic axioms, (3.2) monotonicity axioms, (3.3) consistency, (3.4) proportionality and representation axioms (3.5) axioms related to strategic voting, and, lastly, we consider some (3.6) computational aspects.

Since both procedures from the previous section are defined in terms of existing ABC rules, instead of investigating whether a procedure $\mathcal{R}^\rho$ satisfies an ABC property P , we consider a weaker notion: inheritance. This captures the idea that if all rules in ρ from either two-step procedure $\mathcal{R}^\rho$ satisfy a property P , the procedure itself should satisfy P . Strictly speaking, this is never possible as the rules in ρ are ABC rules and $\mathcal{R}^\rho$ is an ABCC procedure. It is important to note that this section does not contain axioms that are ‘unique’ to the ABCC context. However, it contains axioms ‘lifted’ to the ABCC setting. The ABC setting can be seen as a restricted ABCC setting, where there is only one community consisting of all voters, i.e. $\mathcal{Q} = [n]$. Any ABC rule can also be simulated by both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$. If we take $\rho = (\mathcal{R})$ and apply either procedure to $E = (A, [n], k)$ we are simulating $\mathcal{R}(A, k)$ in the ABCC setting. We now abusively define inheritance.

Definition 3.1. An ABCC procedure $\mathcal{R}^\rho$ *inherits* a property P if whenever all ABC rules $\mathcal{R}_i \in \rho$ satisfy P , $\mathcal{R}^\rho$ also satisfies P for any election $E = (A, \mathcal{Q}, k)$.

$\mathcal{R}_i \in \rho$ satisfying P must be understood as $\mathcal{R}^\rho$ with $\rho = (\mathcal{R}_i)$ satisfying P on all elections $E = (A, [n], k)$.

Inheriting a property P is weaker than satisfying a property P , as it is not the case that an ABCC procedure that inherits a property also always satisfies it. The other way around, if a $\mathcal{R}^\rho$ does not inherit a property P , it also immediately does not satisfy P . Furthermore, recall that $\mathcal{R}_{NWP}^\rho$ takes as input a special rule $\mathcal{R}_{s+1}$. Throughout this section this is part of ρ . It is also possible to define the dual of inheritance:

Definition 3.2. A property P is *determined* for $\mathcal{R}^\rho$ if whenever $\mathcal{R}^\rho$ satisfies P all rules $\mathcal{R}_i$ from ρ satisfy P .

Determination captures the idea that the underlying rule determine whether a given property P will hold for a procedure $\mathcal{R}^\rho$. This is a demanding property, that is included primarily as a theoretical counterpart to inheritance. Almost no natural procedure will satisfy this, because the reason for

satisfying an axiom may depend on the community structure and not on the underlying rules. As an example, let $n = 100$, $m = 4$, $k = 2$ and $\mathcal{Q} = [51, 49]$. Let the ballot of the first 51 voters be $\{a, b\}$ and of the other 49 $\{c, d\}$. Under $\mathcal{R}_{RSP}^\rho$ with ρ consisting of all $\mathcal{R}_{AV}$'s, the outputs satisfies EJR, but $\mathcal{R}_{AV}$ does not.

Finally, diversity and all criteria from Definition 1.4 are assumed to hold for all weights for $\mathcal{R}_{RSP}^\rho$.

3.1 Basic axioms

The first part of this section contains a discussion of two basic axioms: *neutrality* and *anonymity*. As these axioms are not the deepest and most interesting, the goal of discussing them will be to get acquainted with the definitions of inheritance and ABCC properties. The following definitions are based on a paper by Ozkes and Sanver (2021). For a further discussion of anonymity, neutrality and their interplay with other basic axioms, the reader is referred to the same source.

First, we define a *permutation* on a set X to be a bijection $\sigma : X \rightarrow X$. Define $A^\sigma = (A(\sigma(1)), A(\sigma(2)), \dots, A(\sigma(n)))$ to be the profile where the names of the voters have been permuted by a $\sigma : N \rightarrow N$. The interpretation of the following axiom is that all voters should be treated equally.

Axiom 2 (Anonymity). Let $E = (A, \mathcal{Q}, k)$ be an ABCC election. An ABCC procedure $\mathcal{R}^\rho$ is *anonymous* if and only if $\mathcal{R}^\rho(A, \mathcal{Q}, k) = \mathcal{R}^\rho(A^\sigma, \mathcal{Q}, k)$ for all permutations $\sigma : N \rightarrow N$.

It is fairly straightforward to see that this axiom will not be satisfied in the ABCC context, as voters are not treated ‘equally’ depending on the community they belong to. The following axiom, neutrality, captures a kind of fairness towards candidates. Define $A_\sigma = (\sigma(A(1)), \sigma(A(2)), \dots, \sigma(A(n)))$ as the profile where a permutation $\sigma : C \rightarrow C$ has been applied to each ballot.

Axiom 3 (Neutrality). Let $E = (A, \mathcal{Q}, k)$ be an ABCC election. An ABCC procedure $\mathcal{R}^\rho$ is *neutral* if and only if $\sigma(\mathcal{R}^\rho(A, \mathcal{Q}, k)) = \mathcal{R}^\rho(A_\sigma, \mathcal{Q}, k)$.

We can now proceed to show that anonymity is not inherited.

Proposition 3.1. *Neither $\mathcal{R}_{RSP}^\rho$ nor $\mathcal{R}_{NWP}^\rho$ inherits anonymity.*

Proof. Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = [2, 2]$. Voters 1 and 2 are part of N_1 , and 3 and 4 are part of N_2 . We define a permutation $2 \mapsto 3$, and the identity function on the rest. This leads to the following two profiles. The left profile is the original one, the profile on the right is the permuted profile.

1 : $\{a, b\}$	1 : $\{a, b\}$
2 : $\{c, d\}$	2 : $\{a, b\}$
3 : $\{a, b\}$	3 : $\{c, d\}$
4 : $\{c, d\}$	4 : $\{c, d\}$

For $\mathcal{R}_{RSP}^\rho$ take $\rho = (\mathcal{R}_{AV}, \mathcal{R}_{AV})$. Note that technically speaking we take $\mathcal{R}_{RSP}^\rho$ with $\rho = (\mathcal{R}_{AV})$ on $E_i = (A_i, [n_i], q_i)$. The original output is $\binom{C}{2}$. The output on the permuted profile is $\{a, b\} \times \{c, d\}$. For $\mathcal{R}_{NWP}^\rho$ take $\rho = (\mathcal{R}_{PAV}, \mathcal{R}_{PAV}, \mathcal{R}_{AV})$. The original output is $\{a, b\} \times \{c, d\}$, but the permuted output is $\binom{C}{2}$. $\square$

Proving neutrality is omitted, as it is not a property that is relevant to the main discussion of this thesis nor does the proof strategy reappear in the remainder of this thesis. Furthermore, it is fairly immediate to see that it is inherited. I support this claim by an example.

Example 3.1. Let $n = 8$, $m = 4$, $k = 3$ and $\mathcal{Q} = [5, 3]$. Note that $\mathbf{q} = (2, 1)$. The profile on the left is the original profile and the profile on the right is the permuted profile.

$3 \times \{a, b\}$	$3 \times \{\sigma a, \sigma b\}$
$2 \times \{a, c\}$	$2 \times \{\sigma a, \sigma c\}$
$2 \times \{a, d\}$	$2 \times \{\sigma a, \sigma d\}$
$1 \times \{a\}$	$1 \times \{\sigma a\}$

Note that $\mathcal{R}_{AV}$ is neutral. Suppose that ρ consists of all $\mathcal{R}_{AV}$'s in both cases.

For $\mathcal{R}_{RSP}^\rho$, note that $\{a, b\}$ is the winner in the first community and $\{a\}$ is the winner in the second community. The winner sets are contained in $\{\{a, b, c\}, \{a, b, d\}\}$ where c and d are extensions as the original W 's were smaller than k . Now, it does not matter whether we apply σ to this final winner set or we run $\mathcal{R}_{RSP}^\rho$ on the permuted profiles, as all underlying rules are neutral.

We see the same for $\mathcal{R}_{NWP}^\rho$. For this rule we would have to rerun the following elections:

$5 \times \{a, b, c\}$	$3 \times \{\sigma a, \sigma b, \sigma c\}$
$2 \times \{a, d, x\}$	$2 \times \{\sigma a, \sigma d, \sigma x\}$

Here $x \in \{b, c\}$. The winners in both cases are $\{\{a, b, c\}, \{a, b, d\}\}$ and $\{\{\sigma a, \sigma b, \sigma c\}, \{\sigma a, \sigma b, \sigma d\}\}$. It is again clear that applying a neutral rule to a permuted profile is the same as permuting the winner set of the same neutral rule. $\diamond$

A more relevant basic condition one would like any ABC voting rule to satisfy is *weak efficiency*. It captures the idea that non-approved candidates should not make it into the final winner set at the expense of an approved candidate. The formulation here is based on a paper by Kluiving et al. (2020) who base it on the PhD thesis of Peters (2019).

Axiom 4. An ABCC procedure $\mathcal{R}^\rho$ satisfies *weak efficiency* if for any election $E = (A, \mathcal{Q}, k)$ it is the case that, if $|\bigcup_{i \in N} A(i)| \geq k$ then $W \subseteq \bigcup_{i \in N} A(i)$ for all $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$.

Note that all scoring voting rules — these include AV, PAV, CC — satisfy weak efficiency. The idea behind this is that any candidate with a non-zero score will be preferred by such a method (Lackner and Skowron, 2020a). Both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$ do not inherit this basic axiom, which is a problematic fact.

Proposition 3.2. *Neither $\mathcal{R}_{RSP}^\rho$ nor $\mathcal{R}_{NWP}^\rho$ inherits weak efficiency.*

Proof. Note that $\mathcal{R}_{AV}$ satisfies weak efficiency. For both procedures, let ρ consist of $\mathcal{R}_{AV}$'s only and pick $n = 3$, $m = 4$ and $\mathcal{Q} = [2, 1]$. Consider the following profile:

$2 \times \{a\}$
$1 \times \{b, c\}$

We first show the result for $\mathcal{R}_{RSP}^\rho$. Suppose that $k = 2$. $\mathbf{q} = (2, 1)$, by weak proportionality. In the first community, all subsets of $\binom{C}{2}$ containing a are the winners. $\{\{b\}, \{c\}\}$ contains the winners for the second community. This means that $\{a, d, b\}$ and $\{a, d, c\}$ are one of the final winner sets. But, there exists no $i \in N$ such that $d \in A(i)$ whereas there are voters approving c , for the first winner set, and b for the second winner set.

For $\mathcal{R}_{NWP}^\rho$ pick $k = 2$ pick $k = 3$ We first show that $\mathcal{R}_{RSP}^\rho$ does not inherit weak efficiency. The reasoning for $\mathcal{R}_{NWP}^\rho$ is similar. After running the two community elections, one of the profiles we apply $\mathcal{R}_{s+1} = \mathcal{R}_{AV}$ to is:

$$\frac{2 \times \{a, d\}}{1 \times \{b, c\}}$$

So, $\{a, d\} \in W$. Again, no one approves of d whereas there are voters approving c and b . $\square$

The two procedures not satisfying this axiom is detrimental for their normative appeal. There is no argument to be made for not satisfying weak efficiency from the perspective of descriptive representation, as not satisfying this basic axioms means that in some cases there is no representation at all. Fortunately, it is possible to make a small adjustment to both procedures and make it satisfy weak efficiency. Instead of adding *all* candidates from C to a winning set W_i that is smaller than q_i , we first add candidates from $\bigcup_{j \in N \setminus N_i} A(j)$ whenever we have to break ties to extend winner sets that are too small and then, if the set is still too small, we add candidates from $C \setminus \bigcup_{j \in N \setminus N_i} A(j)$. This hinders any non-approved candidates from entering the community winner sets and the final winner set. A formal proof is omitted, as the result is fairly straightforward. The idea is that no community winner set will include a non-approved candidate over an approved one. Tie-breaking, then, is the only way this can happen. The redefinition prevents this. This leads to:

Observation 3.3. *The redefined versions of $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ inherit weak efficiency.*

Now, we consider *Pareto efficiency*. The formulation is from Lackner and Skowron (2020a). This captures the idea that a committee that is *dominated* by another committee, should not be chosen by our procedure. A committee W_1 dominates another committee W_2 if (1) every voter has at least as many approved candidates in W_1 as in W_2 and (2) at least one candidate strictly approves of more candidates in W_1 . Formally:

Definition 3.3. Let $E = (A, \mathcal{Q}, k)$ be an election. A committee W_1 *dominates* a committee W_2 if:

1. For $i \in N$ it is the case that $|A(i) \cap W_1| \geq |A(i) \cap W_2|$.
2. There exists a $j \in N$ with $|A(j) \cap W_1| > |A(j) \cap W_2|$.

Axiom 5. An ABCC procedure $\mathcal{R}^\rho$ satisfies *strong Pareto efficiency* if $\mathcal{R}^\rho$ never outputs dominated committees. $\mathcal{R}^\rho$ satisfies *weak Pareto efficiency* if for all election instances $(A, \mathcal{Q}, k)$ it holds that if $W_2 \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$ and W_1 dominates W_2 then $W_1 \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$.

It follows from the definition that strong Pareto efficiency implies the weak variant. $\mathcal{R}_{AV}$ satisfies strong Pareto efficiency and $\mathcal{R}_{CC}$ satisfies weak Pareto efficiency. Proofs of these facts can be found in Lackner and Skowron (2020b). Neither procedure we defined in the previous chapter satisfy this axiom. The proof of the following proposition is based on an example from Lackner and Skowron (2020b).

Proposition 3.4. $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ do not inherit weak (or strong) Pareto efficiency.

Proof. First, consider $\mathcal{R}_{RSP}^\rho$. Let $n = 24$, $m = 4$, $k = 2$ and $\mathcal{Q} = [12, 12]$. By diversity, $\mathbf{q} = (1, 1)$. Consider this profile:

$$\begin{array}{r}
10 \times \{a, c\} \\
1 \times \{b\} \\
1 \times \{b, c\} \\
\hline
10 \times \{a, d\} \\
1 \times \{b, d\} \\
1 \times \{b\}
\end{array}$$

Pick $\rho = (\mathcal{R}_{AV}, \mathcal{R}_{AV})$. Then, c wins in the first community and d wins in the second community. Hence, $\{c, d\}$ is the winner set. $\{a, b\}$, however dominates $\{c, d\}$.

Now, we consider $\mathcal{R}_{NWP}^\rho$. With $\rho = (\mathcal{R}_{AV}, \mathcal{R}_{AV}, \mathcal{R}_{AV})$. We use the same counterexample as above, except for the size of the winner set, which is changed to $k = 1$. Remark that $\{c\}$ is the winner in the first community and $\{d\}$ in the second community. Since both communities have equal size, in the end $\{c\}$ and $\{d\}$ are the winning coalitions. $\{a\}$, however, dominates both sets, since it is approved of by 20 voters and both $\{c\}$ and $\{d\}$ only by 11, so in both cases there are 9 voters approving of strictly more candidates. $\square$

Pareto efficiency can be seen as a basic requirement for any *utilitarian rule*. This class of rules selects candidates that enjoy the largest number of support from voters. In the context of ABC(C) voting, this means that a rule should maximise the number of approvals received by the winning candidates (Lackner and Skowron, 2020b). Utilitarian rule are contrasted by *representative* rules, which try to maximise the voters with at least one approved candidate in the final winner set. It is intuitively clear that these two desiderata — maximising total happiness and representing as many people as possible — are incompatible in many contexts. Since the two procedures $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ have been implemented with the aims of representing groups of voters, it is not a big downside that they do not inherit Pareto efficiency.

3.2 Monotonicity

This section contains a discussion of two axioms: committee monotonicity and support monotonicity. Committee monotonicity captures the idea that if we change the size of winning committee from k to $k + 1$, the k winners should be contained in the winning set of size $k + 1$. More practically, if we first vote for a president and then for a president and a prime minister, then, the winning president from the first election should also win in the second election. Support monotonicity captures the basic idea that adding votes for a given (set of) candidate(s) should not harm how well they perform in the election. These two axioms are measures of fairness towards candidates.

The following formulation of committee monotonicity for irresolute rules is based on the paper by Elkind et al. (2017).

Axiom 6. A procedure $\mathcal{R}^\rho$ satisfies *committee monotonicity* if for any ABCC election $E = (A, \mathcal{Q}, k)$ the following conditions hold:

1. If $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$, then there exists a $W' \in \mathcal{R}^\rho(A, \mathcal{Q}, k + 1)$ such that $W \subseteq W'$.
2. If $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k + 1)$, then there exists a $W' \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$ such that $W' \subseteq W$.

This axiom is only inherited by one of the two procedures. The inclusion of house monotonicity is not a coincidence, as committee monotonicity and house monotonicity are related notions. For more details, see 4.1 of the book by Lackner and Skowron (2023).

Theorem 3.5. $\mathcal{R}_{RSP}^\rho$ inherits committee monotonicity if the weights $\mathbf{q} = (q_1, \dots, q_s)$ are determined by an apportionment rule that satisfies house monotonicity.

Proof. This proof is by a case distinction, based on lines 15-23 from the algorithm formulation of $\mathcal{R}_{RSP}^\rho$ (2). Recall that after the winning coalitions for all communities have been returned, the final winning coalition is too small, i.e., $|W| < k$. This happens when, for instance, all community winning coalitions for a community i have overlap with all winning coalitions of community j . For an example of this, the reader is referred to Example 2.8.

If it is indeed the case that $|W| < k$, recall that the set W is extended in the following way. For all elements of $\binom{C \setminus W}{|W| - k}$, take the union with W and return these as the final winning sets. By construction, this union is of size k . Recall that the elements of $\binom{C \setminus W}{|W| - k}$ are called ‘extensions’ in this context.

The case distinction is given as follows: (i) the final winning coalitions are of size k and we do not have to extend and (ii) the final winning coalitions are not of size k and we have to extend.

(i) First, we prove condition 1 and then 2.

1. Take any $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. Consider $\mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k+1)$. By house monotonicity of the determination of the weights $\mathbf{q}$, the vector of weights for the new election is given by: $\mathbf{q} = (q_1, \dots, q_\ell + 1, \dots, q_s)$ for an $\ell \in [s]$. Note that $W = W_1 \cup \dots \cup W_s$ and $|W| = k$. For $W_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$, we know that, by monotonicity of all rules in ρ , there exists a $W'_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell + 1)$ such that $W_\ell \subseteq W'_\ell$. All other community elections return the same winner sets. This means that $W' = W_1 \cup \dots \cup W'_\ell \cup \dots \cup W_s \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k+1)$ is a set such that $W \subseteq W'$. The reasoning being this is, is that for all $i \neq \ell$, $W_i \subseteq W'$ and for W_ℓ , it is the case that $W_\ell \subseteq W'_\ell \subseteq W'$.
2. For condition 2, take $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k+1)$. The reasoning is similar to the proof of condition 1. By house monotonicity, there is one community ℓ such that their weight decreases from q_ℓ to $q_\ell - 1$ if we consider $\mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. Now, note that $W = W_1 \cup \dots \cup W_s$. By committee monotonicity of all rules in ρ , we know that for $W_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$ there exists $W'_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell - 1)$. Now, as all other community elections return the same winner sets, $W' = W_1 \cup \dots \cup W'_\ell \cup \dots \cup W_s$ is the set that is included in W .

(ii) Now, we consider the case in which the final winning set W is too small and has to be extended to a winning set of size k .

1. Take any $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. Note that W can be written as $W_1 \cup \dots \cup W_s \cup X$, where W_i for all $i \in [s]$ refers to the community winner sets and X is the extension from $\binom{C \setminus \bigcup_{i \in [s]} W_i}{k - |W_1 \cup \dots \cup W_s|}$. Note that the size of X is $k - |W_1 \cup \dots \cup W_s|$. Consider $\mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k+1)$. As in case (i), the vector of weights for the new election is given by: $\mathbf{q} = (q_1, \dots, q_\ell + 1, \dots, q_s)$.

It could be the case that all $W'_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell + 1)$ always include a candidate $x \in X$, i.e., for all such W'_ℓ it is the case that they can be written as $W_\ell \cup \{x\}$. If this is the case, $|W_1 \cup \dots \cup W'_\ell \cup \dots \cup W_s| = |W_1 \cup \dots \cup W_s| + 1 = |W| + 1$, as there is now one candidate $x \in W'_\ell$ that does not appear in $W_1 \cup \dots \cup W_s$. If this is the case, we have to extend $W_1 \cup \dots \cup W'_\ell \cup \dots \cup W_s$ by $(k+1) - (|W_1 \cup \dots \cup W_s| + 1) = k - |W_1 \cup \dots \cup W_s|$ candidates. Note that this means that this extension is of the same cardinality as X , the original extension for $\mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. This means that we can pick $W' = W_1 \cup \dots \cup W'_\ell \cup \dots \cup W_s \cup X \setminus \{x\} \cup \{y\}$,

where y is any unchosen candidate. It is clear that $W \subseteq W'$, as all $W_i \subseteq W_i$, $W_\ell \subseteq W'_\ell = W_\ell \cup \{x\}$ and $X \subseteq X \setminus \{x\} \cup \{x\}$.

The other case, where $W'_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$ does not always include a candidate $x \in X$, we can take $W' = W_1 \cup \dots \cup W_s \cup X \cup \{y\}$, where $y \in C$. In this case, the size of $W_1 \cup \dots \cup W_s$ does not increase and we extend with a bigger set $X \cup \{y\}$.

2. Take $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k+1)$. Again, $W = W_1 \cup \dots \cup W_s \cup X$, where X is the extension. Suppose that the weight of community ℓ decreases from q_ℓ to $q_\ell - 1$. Their winner set from W , W_ℓ , loses a candidate. We have to show that there exists a set $W' = W'_1 \cup \dots \cup W'_s \cup X'$ such that $W' \subseteq W$. Do note that we know for sure that for all $i \neq \ell$ it is the case that the elections are the same. So, we can take $W'_i = W_i$ for all $i \in [s] \setminus \{\ell\}$. $W'_\ell \subseteq W_\ell$, by committee monotonicity of $\mathcal{R}_\ell$. If we can show that the extension X' of $\bigcup_{i \neq \ell} W_i \cup W'_\ell$ is included in X we are done.

Note that $W_\ell \setminus \{c\} = W'_\ell$. If it is the case that $c \in W_i$ for $i \neq \ell$ it is the case that $\bigcup_{i \neq \ell} W_i \cup W'_\ell = W$, hence we take $X' = X \setminus \{x\}$ for any $x \in X$. Otherwise, if $c \notin W_i$ for $i \neq \ell$, the size of $\bigcup_{i \neq \ell} W_i \cup W'_\ell$ decreases by 1. This means that X' has to extend by picking a set of size $(k-1) - (|\bigcup_{i \neq \ell} W_i \cup W'_\ell| - 1) = k - (|\bigcup_{i \neq \ell} W_i \cup W'_\ell|)$, which is exactly the size of X . So, we may pick $X' = X$.

As these cases are exhaustive, this completes the proof. $\square$

The other procedure, $\mathcal{R}_{NWP}^\rho$, does not inherit this property. For this we use that $\mathcal{R}_{AV}$ satisfies committee monotonicity, which is proven by Sánchez-Fernández and Fisteus (2018).

Proposition 3.6. $\mathcal{R}_{NWP}^\rho$ does not inherit committee monotonicity.

Proof. Note that $\mathcal{R}_{AV}$ satisfies committee monotonicity. Let ρ consist of all $\mathcal{R}_{AV}$'s. Furthermore, let $n = 6$, $m = 4$ and $\mathcal{Q} = [3, 3]$. Now consider the following profile:

$$\begin{array}{c} 2 \times \{a\} \\ 1 \times \{b\} \\ \hline 2 \times \{c\} \\ 1 \times \{d\} \end{array}$$

For $k = 1$, we know $\mathcal{R}_{AV}(N_1, 1) = \{a\}$ and $\mathcal{R}_{AV}(N_2, 1) = \{c\}$. $\mathcal{R}_{NWP}^\rho$ now applies $\mathcal{R}_{AV}$ to the following profile to obtain the final winner set.

$$\begin{array}{c} 3 \times \{a\} \\ 3 \times \{c\} \end{array}$$

Hence, $W = \{\{a\}, \{c\}\}$. For $k = 2$, observe that $\mathcal{R}_{AV}(N_1, 2) = \{a, b\}$ and $\mathcal{R}_{AV}(N_2, 2) = \{c, d\}$. One of the sets, $\mathcal{R}^\rho$ applies $\mathcal{R}_{AV}$ to is:

$$\begin{array}{c} 3 \times \{a, b\} \\ 3 \times \{c, d\} \end{array}$$

In this case, all candidates get three votes. Hence, $\mathcal{R}(A, k+1) = \binom{C}{2}$. Remark that this includes $W' = \{b, d\}$. So, there exists a set $W \in \mathcal{R}(A, k+1)$ but there exists no $W' \in \mathcal{R}(A, k)$ such that $W' \subseteq W$. $\square$

We now consider support monotonicity. This axiom can be divided into two variants. We first do *support monotonicity without additional voters*. The other variant comes *with* additional voters. The following definition is from Sánchez-Fernández and Fisteus (2018) and the formulation is from Lackner and Skowron (2023). A short notational remark: A_{i+X} refers to profile A where voter i adds X to her ballot. I will write A_ℓ^{i+X} instead of $A_{\ell_{i+X}}$ for community elections.

Axiom 7. A $\mathcal{R}^\rho$ satisfies *support monotonicity without additional voters* if for every election instance $(A, \mathcal{Q}, k)$, $i \in N$ and a candidate set $X \subseteq W$ it holds that

1. if $X \subseteq W$ for all $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$ then $X \subseteq W'$ for all $W' \in \mathcal{R}^\rho(A_{i+X}, \mathcal{Q}, k)$ and
2. if $X \subseteq W$ for some $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$ then $X \subseteq W'$ for some $W' \in \mathcal{R}^\rho(A_{i+X}, \mathcal{Q}, k)$

This axiom is inherited by both procedures. In the proofs, I use ‘extension’ in the same sense as in the proof of Theorem 3.5.

Theorem 3.7. $\mathcal{R}_{RSP}^\rho$ inherits support monotonicity without additional voters.

Proof. Let $E = (A, \mathcal{Q}, k)$ be an election. Suppose that ρ consists of rules that all satisfy support monotonicity. The second condition is more straightforward than the first condition, so we consider it first and then build on the proof strategy for the first condition. Fix voter i and let her be a member of community ℓ . We consider two cases:

2. Suppose that $X \subseteq W$ for some $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. Note that $W = W_1 \cup \dots \cup W_s \cup Y$ where Y is a possible extension. Two cases are distinguished:
 - (a) Suppose that $X \cap W_\ell = \emptyset$. This means that $X \subseteq W \setminus W_\ell$. Note that there is at least one winner set W' such that $W \setminus W_\ell \subseteq W'$, because for elections $\mathcal{R}_j(A_j^{i+X}, q_j)$ where $j \neq \ell$ it holds that $\mathcal{R}_j(A_j^{i+X}, q_j) = \mathcal{R}_j(A_j, q_j)$, i.e., ‘nothing changes’ if we add X to i ’s ballot. Hence, there is at least one $W' \in \mathcal{R}_{RSP}^\rho(A_{i+X}, \mathcal{Q}, k)$ such that $X \subseteq W'$. This W' is of the following form: $W' = \bigcup_{j \neq \ell} W_j \cup W'_\ell \cup Y$, where W'_ℓ is any element of $\mathcal{R}_\ell(A_\ell^{i+X}, q_\ell)$
 - (b) Suppose that $X \cap W_\ell \neq \emptyset$. Write $X = X_1 \cup \dots \cup X_s \cup Y'$, where $X_i = X \cap W_i$ for $i \in [s]$ and $Y' = X \cap Y$. Note that $X_\ell \subseteq W_\ell$ for some $W_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$. By support monotonicity, it holds that $X_\ell \subseteq W'_\ell$ for some $W'_\ell \in \mathcal{R}_\ell(A_\ell^{i+X}, q_\ell)$. For $\mathcal{R}_{RSP}^\rho(A_{i+X}, \mathcal{Q}, k)$, we can now take $W' = \bigcup_{i \neq \ell} W_i \cup W'_\ell \cup Y$. Note that all W_i ’s where $i \neq \ell$ appear in all $\mathcal{R}_j(A_j^{i+X}, q_j)$, as $A_j^{i+X} = A_j$ because voter i is not a member of community j .
1. Suppose that $X \subseteq W$ for all $W \in \mathcal{R}_{RSP}^\rho(A, \mathcal{Q}, k)$. Again, write $W = W_1 \cup \dots \cup W_s \cup Y$. We do another case distinction:
 - (a) Suppose that (1) $\bigcap \mathcal{R}_i(A_i, q_i) = \emptyset$ for all $i \in [s]$ and (2) $\bigcup \mathcal{R}_i(A_i, q_i) \cap \bigcup \mathcal{R}_i(A, q_j) = \emptyset$ for all $i \neq j$. Informally, the first condition entails that there is more than one winner set for each community, and each of these winner sets has different candidates. The second condition captures the idea that each community has different ‘representatives’. Note that in this case the condition is satisfied vacuously, as there can be no fixed set X in all final winner sets. To see this, consider any two distinct final winner sets $W = W_1 \cup \dots \cup W_s \cup Y$ and $W' = W'_1 \cup \dots \cup W'_s \cup Y'$. The first condition implies that for all $i \in [s]$, W_i and W'_i are disjoint. The second condition implies that for all $i \neq j$ W_i and W_j are disjoint. Hence, W and W' are disjoint. So, there can be no set X contained in both W and W' .

- (b) Suppose that (1) $\bigcap \mathcal{R}_i(A_i, q_i) = \emptyset$ for all $i \in [s]$ and (2) $\bigcup \mathcal{R}_i(A_i, q_i) \cap \bigcup \mathcal{R}_i(A, jq_j) \neq \emptyset$ for some $i \neq j$. First, note that we never have to extend in this case, because of the second condition. This condition tells us that each community has ‘unique’ representatives. If we take the union over the community winner sets, there is no overlap and we never extend.

Now, take any two final winner sets $W = W_1 \cup \dots \cup W_s \cup Y$ and $W' = W'_1 \cup \dots \cup W'_s \cup Y'$. By the first condition it holds that $W_i \cap W'_i$ and $W_j \cap W'_j$ are empty. According to the second condition, it is the case that there is some W_i and some W_j such that $W_i \cap W_j \neq \emptyset$. The only way that W and W' have a shared winner set is if $(W_i \cap W_j) \cup (W'_i \cap W'_j) \subseteq X$ is non-empty.

To finish this case, consider a voter that adds X to her ballot. If she is not in community i or j , X reappears in all final winner sets trivially. If she is in either one of the two communities and she adds X to her ballot, X will reappear in some winner set again, by the second condition of support monotonicity. Hence, X reappears in all final winner sets.

- (c) Suppose that (1) $\bigcap \mathcal{R}_i(A_i, q_i) \neq \emptyset$ for all $i \in [s]$ and (2) $\bigcup \mathcal{R}_i(A_i, q_i) \cap \bigcup \mathcal{R}_i(A, jq_j) = \emptyset$ for all $i \neq j$. As the community winner sets are not disjoint, it might be the case that we have to extend the final winner sets. If we have to extend, however, it can never be the case that X intersects with the extension, except if $k = |C|$. Recall that we pick extensions from $\binom{C \setminus W}{k - |W|}$. The only way to always pick one candidate is if $|C \setminus W| = |C| - |W| = k - |W|$ for all W . Note that this is the case only if $|C| = k$. If this is the case, however, the winner set is trivial: $W = C$. This means that we again can assume that we do not have to extend.

Now, take any two final winner sets $W = W_1 \cup \dots \cup W_s \cup Y$ and $W' = W'_1 \cup \dots \cup W'_s \cup Y'$. By condition (2), it is the case that $W_i \cap W_j$ and $W'_i \cap W'_j$ are empty for all $i \neq j$. The only way a set X can reappear in all final winner sets is if it reappears in all community winner sets, for some community. If a voter adds X to her ballot and she is part of such a community, X reappears in all those community winner sets, by support monotonicity. Hence, it reappears in all final winner sets. If this voter is not part of such a community ‘contributing’ to X , X reappears irregardless of what that voter does.

- (d) Suppose that (1) $\bigcap \mathcal{R}_i(A_i, q_i) \neq \emptyset$ for all $i \in [s]$ and (2) $\bigcup \mathcal{R}_i(A_i, q_i) \cap \bigcup \mathcal{R}_i(A, jq_j) \neq \emptyset$ for some $i \neq j$. We do not have to extend, by the same reasoning as for (c). The reasoning for this case is a combination of cases (b) and (c). The set X that reappears in all community winner sets has a part that reappears in all final winner sets of one community, by condition (1). If the voter that adds X to her ballot is part of this community, X again reappears in all final winner sets, because of support monotonicity of the community voting rule.

For the part of X that is not ‘contributed’ by the community with non-disjoint winner sets, consider the following reasoning: take any two final winner sets $W = W_1 \cup \dots \cup W_s \cup Y$ and $W' = W'_1 \cup \dots \cup W'_s \cup Y'$. Without loss of generality we suppose that W_1 and W'_1 are non-disjoint, but we also assume that the voter that adds X to her ballot is not part of community 1. Because of condition (2) it could be the case that a subset of X , for which we use X' , reappears in both winner sets W and W' because there is some W_i and some W_j such that $W_i \cap W_j \neq \emptyset$. The only way that W and W' have a shared winner set is if $(W_i \cap W_j) \cup (W'_i \cap W'_j) \subseteq X$ is non-empty. Now we are in the situation of case (b) and the final reasoning is similar to the last paragraph of that case.

As these cases are exhaustive, this finishes the proof. $\square$

The rough idea of the proof above is that if voter i supports a set X in a community ℓ where the community rule $\mathcal{R}_\ell$ satisfies support monotonicity, it is case that the part of X , X_ℓ , that was already in the winner set of a community ℓ , will remain or be accompanied by other candidates from X . If it is not the case that she is part of a community supporting X , the other communities' winner sets on their own already guarantee that X reappears in some or all winner sets. This idea is fairly straightforward and the proof for $\mathcal{R}_{NWP}^\rho$ follows it as well. Recall that for $\mathcal{R}_{NWP}^\rho$, we construct profiles based on the outcomes of the community elections, the set of all these profiles was denoted by $\mathbf{A}$.

Theorem 3.8. $\mathcal{R}_{NWP}^\rho$ inherits support monotonicity without additional voters.

Proof. We first check condition 1. Take X such that $X \subseteq W$ for all $W \in \mathcal{R}_{NWP}^\rho(A, \mathcal{Q}, k)$. This means that for any profile $\Pi \in \mathbf{A}$, it is the case that for all $W \in \mathcal{R}_{s+1}(\Pi, k)$, $X \subseteq W$. Suppose that voter i , who is part of community ℓ , starts supporting X . Hence, we consider $\mathcal{R}_{NWP}^\rho(A_{i+X}, \mathcal{Q}, k)$.

As the assumption that $W \in \mathcal{R}_{NWP}^\rho(A, \mathcal{Q}, k)$ says nothing about the community profile, we do a case distinction. Suppose that $X \not\subseteq W_\ell$ for some $W_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$. An increase in support for X , might lead to X being part of more winning sets. If this is the case, there are more elections $\Pi \in \mathbf{A}$ where the support for X is increased. Since $\mathcal{R}_{s+1}$ is support monotone, it is the case that $X \subseteq W$ for all $W \in \mathcal{R}_{NWP}^\rho(A_{i+X}, \mathcal{Q}, k)$.

If $X \subseteq W_\ell$ for all $W_\ell \in \mathcal{R}_\ell(A_\ell, q_\ell)$, it is also the case that $X \subseteq W_\ell$ if X is added to i 's ballot. In this case, all $\Pi \in \mathbf{A}$ remain the same and it is again the case that $X \subseteq W$ for all $W \in \mathcal{R}_{NWP}^\rho(A_{i+X}, \mathcal{Q}, k)$.

We now check condition 2. Take X such that $X \subseteq W$ for all $W \in \mathcal{R}_{NWP}^\rho(A, \mathcal{Q}, k)$. This means that there exists a profile $\Pi \in \mathbf{A}$, such that there is a $W \in \mathcal{R}_{s+1}(\Pi, k)$ where $X \subseteq W$. We do the same case distinction as for condition 1 and obtain the same outcome.

If X is not part of any of the community winning sets, i adding X to her ballot might cause X to become part of a community winning set. This could change at least one Π in $\mathbf{A}$ in such a way that $|N_\ell|$ voter will now approve X . In particular, this means that we rerun a $\mathcal{R}_{s+1}(\Pi_{i+X}, k)$ and, for all other $\Pi' \in \mathbf{A}$, we do $\mathcal{R}_{s+1}(\Pi', k)$. Because of support monotonicity, at least one of the winner sets of these elections will include X .

Otherwise, if X is part of a community winner set, the result follows directly from support monotonicity. As in the proof of condition 1. $\square$

Recall that we mentioned that there are two variants of support monotonicity. The other variant of support monotonicity can be formulated as follows:

Axiom 8. $\mathcal{R}^\rho$ inherits support monotonicity with additional voters, if whenever ρ consists of voting rules all satisfying support monotonicity with additional voters it is the case that for every $(A, \mathcal{Q}, k)$, $i \in N$ and $X \subseteq C$ it holds that:

1. if $X \subseteq W$ for all $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$, then $X \subseteq W'$ for all $W' \in \mathcal{R}(A_{N_j+X}, k)$ for $j \leq s$.
2. if $X \subseteq W$ for some $W \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$, then $X \subseteq W'$ for some $W' \in \mathcal{R}(A_{N_j+X}, k)$ for $j \leq s$.

Whether this property is inherited by $\mathcal{R}_{NWP}^\rho$ or $\mathcal{R}_{RSP}^\rho$ is an open question. From a high-level perspective support monotonicity *without* addition voters is inherited, because the final winner sets are made up out of community winner sets. If these are monotonic under added support, the final winner set is too. Changing 'without' to 'with' does not interfere with this intuition. Therefore, I expect this axiom to be inherited by both procedures as well. It is also possible to define *support monotonicity*

with additional communities if one replaces A_{N_j+X} by $A_{Q:X}$. Here, we add a community consisting of one voter that with X in her ballot. Determining whether this property holds if the underlying rules satisfy support monotonicity with additional candidates is non-trivial. It is omitted in this thesis.

3.3 Consistency

Consistency captures the idea that if we combine two disjoint elections with the same winning committee, then the elections combined should also lead to this winning committee. As an example, if two friend groups on different tables decide for the same dish separately, they should also decide for this dish if they come together and order as one group. For a systematic discussion of consistency in the ABC context, the reader is referred to the authoritative work by Lackner and Skowron (2020a). Historically, the first discussion of this axiom was in the context of preference aggregation. Most notably in the work by Smith (1973) and Young (1974, 1975).

Since it is not obvious what ‘combining’ means in the ABCC setting – *is it adding a new community or extending an existing community?* – this axiom will have to be redefined. There are two ‘natural’ ways of defining consistency as ABCC axioms. First, I exemplify these two kinds of consistency. In the following examples, some variables will not be defined explicitly as it is clear from the tables containing the profiles.

Example 3.2. Consider the following two profiles, both with $m = 4$.

$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \end{array}$	$1 \times \{a, b, c, d\}$
---	---------------------------

We can either extend Q and add the two communities as new and separate communities or we add together two communities.

$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \\ \hline 1 \times \{a, b, c, d\} \end{array}$	$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ 1 \times \{a, b, c, d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \end{array}$
---	--

On the left, we add a third community to the election. On the right, we assume that the person approving $\{a, b, c, d\}$ belongs to first community. These are two kinds of consistency that arise naturally. The left extension is similar to regular consistency, since we ‘paste’ a new profile next to an old one. If the outcomes of the procedures overlapped to begin with, the procedure should now output the intersection. The right extension requires more care. Here we want the conjoined profile to have the same winner under $\mathcal{R}^\rho$ if all the communities that are conjoined have some internal overlap and the $\mathcal{R}^\rho$ itself had some overlap. $\diamond$

The first kind of consistency is an *external* one, since we add another community. The other I will call *internal*. Unfortunately, the discussion of this variant will be omitted. First of all, because it is not important for the main story of this thesis, and, secondly, because it is not straightforward to define it.

I will write $\mathcal{Q} + \mathcal{Q}'$ if we add together two community structures, remark that this is possible only if we add together the ballots of the two elections. The formulation is based on the definition from Lackner and Skowron (2020a).

Axiom 9 (External consistency). A procedure $\mathcal{R}^\rho$ satisfies *external consistency* if for every $k \geq 1$ and two profiles $A : N \rightarrow \wp(C)$ and $A' : N' \rightarrow \wp(C)$ with $N \cap N' = \emptyset$, if $\mathcal{R}^\rho(A, \mathcal{Q}, k) \cap \mathcal{R}^\rho(A', \mathcal{Q}', k) \neq \emptyset$ and $\max\{\mathcal{Q}, \mathcal{Q}'\} < \frac{k}{2}$, then $\mathcal{R}^\rho(A + A', \mathcal{Q} + \mathcal{Q}', k) = \mathcal{R}^\rho(A, \mathcal{Q}, k) \cap \mathcal{R}^\rho(A', \mathcal{Q}', k)$.

I now present the results for these axioms. Note that $\mathcal{R}_{AV}$ satisfies consistency, Lackner and Skowron (2020a) provide a proof of this fact.

Proposition 3.9. $\mathcal{R}_{RSP}^\rho$ does not inherit external consistency from consistent rules.

Proof. Let ρ consist of $\mathcal{R}_{AV}$'s only. Consider the following profiles:

$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \end{array}$		$1 \times \{a, b, c, d\}$
---	--	---------------------------

Pick $k = 3$ and $m = 4$ in both cases. The weights for the election on the left are $\mathbf{q} = (2, 1)$, by the assumptions on the weights. Note that the winner for the election for this election is $\{a, b, d\}$ and for the election on the right the winners are contained in $\binom{C}{3}$. Hence, their intersection is non-empty and equal to $\{a, b, d\}$

We now combine the profile externally:

$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \\ \hline 1 \times \{a, b, c, d\} \end{array}$

The weights are given by $(1, 1, 1)$, because of diversity. The winners in the first community are $\{\{a\}, \{b\}\}$, in the second community it is $\{d\}$ and in the third the winner sets are contained in $\{\{a\}, \{b\}, \{c\}, \{d\}\}$. So, the final winners are $\{\{a, d, b\}, \{a, d, c\}, \{b, d, c\}\}$. This is not the same as $\{a, b, d\}$. $\square$

The other basic procedure also does not satisfy this axiom.

Proposition 3.10. $\mathcal{R}_{NWP}^\rho$ does not inherit external consistency from consistent rules.

Proof. Let ρ consist of $\mathcal{R}_{AV}$'s only. Consider the following profiles:

$\begin{array}{r} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \end{array}$		$100 \times \{a, b, c, d\}$
---	--	-----------------------------

Let $k = 3$ and $m = 4$ in both cases. For the profile on the left, we have to rerun:

$$\begin{array}{c} 5 \times \{a, b, d\} \\ 3 \times \{a, d\} \cup \{x\} \end{array}$$

Note that $x \in \{c, b\}$. In both cases, $\{a, b, d\}$ wins. In the rightmost election, $\binom{C}{3}$ is the output. Now consider the externally combined profile:

$$\begin{array}{c} 3 \times \{a, b\} \\ 2 \times \{d\} \\ \hline 2 \times \{d\} \\ 1 \times \{a\} \\ \hline 100 \times X \end{array}$$

Here $X \in \binom{C}{3}$. It is clear that the 100 votes always ‘dominate’ the other votes. Hence, $X = W$ for all X . Hence, the outcome is $\binom{C}{3}$. But, $\binom{C}{3} \neq \{a, b, d\}$. So, the output of the externally combined elections is not the same as the intersection of the two elections. $\square$

Consistency can be seen as a property that characterises *scoring* voting rules, or, voting rules that maximise a given score function (Lackner and Skowron, 2020a). Examples of this are AV, CC and PAV. The two procedures do not function as scoring mechanisms. This is the case because, in many cases, the community structure inhibits the maximisation of one ‘general’ scoring function.

3.4 Proportionality and representation

This section contains the class of axioms that align most directly with the main topic of this work. First, recall the definition of EJER:

Axiom 10 (EJR restated for ABCC procedures). An ABCC procedure $\mathcal{R}^p$ satisfies *extended justified representation* if for each election instance $E = (A, \mathcal{Q}, k)$, each winning committee $W \in \mathcal{R}^p(E)$ and each ℓ -cohesive group of voters V there exists a voter $i \in V$ with at least ℓ representatives in W ; $\mathcal{A}(i) \cap W \geq \ell$.

Phragmén’s sequential rule and PAV both satisfy this axiom, but these rules give different outputs. We show that this is the case because these two voting rules satisfy different proportionality axioms. We first exemplify their difference. The example is taken from the paper by Peters and Skowron (2020):

Example 3.3. Let $n = 6$, $m = 16$ and $k = 12$. Since we are working in the context of ABC voting, fix $\mathcal{Q} = [6]$. This is the profile:

$$\begin{array}{l} 1 : \{c_1, c_2, c_3, c_4\} \\ 2 : \{c_1, c_2, c_3, c_5\} \\ 3 : \{c_1, c_2, c_3, c_6\} \\ 4 : \{c_7, c_8, c_9\} \\ 5 : \{c_{10}, c_{11}, c_{12}\} \\ 6 : \{c_{13}, c_{14}, c_{15}\} \end{array}$$

Phragmén’s sequential rule picks $W = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_{10}, c_{11}, c_{13}, c_{14}\}$ whereas PAV returns $W = \{c_1, c_2, c_3, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}, c_{13}, c_{14}, c_{15}\}$.

seq-Phragmén treats voters 1, 2 and 3 as a group and assigns them half of the seats, since they form half of the electorate. All other voters are seen as singleton groups and are assigned a sixth of the seats. Phragmén-like rules are proportional with respect to *power*. Meaning that, each fraction of size α , should be given voting power that enables them to decide about α of the final committee. PAV, however, is a type of *welfarist* rule that does not assign each voter an equal amount of power. Instead, it aims to divide *utility* in a proportional way. We see that the group $\{1, 2, 3\}$ gets as much candidates as each singleton $\{4\}$, $\{5\}$ and $\{6\}$ under PAV. This class of rules maximises a welfare function, which does not take into consideration how much power each voter deserves.

In this example, we can see that PAV picks three candidates for every single voter, but it does not take into account that there is a group of three voters that each get the same candidate, whereas there are three individual who are represented by three unique candidates. Intuitively, seq-Phragmén wants each voter to decide about two seats of the final winning committee. The first three voters can combine their power, to elect c_1 , c_2 and c_3 and with their remaining power they elect c_4 , c_5 and c_6 . The remaining voters, 4, 5 and 6 each pick two candidates of their liking. Note that the ABCC model is also designed to assign power to communities that would otherwise not be heard. A big difference, however, is that voters from different communities cannot ‘combine’ their power to elect a shared candidate. $\diamond$

As mentioned in the beginning, Phragmén-like rules satisfy another kind of proportionality than PAV – or any other kind of welfarist rule. This difference can be shown by proving that Phragmén-like rules satisfy *priceability* and *laminar proportionality* (Peters and Skowron, 2020). Without going into detail it is possible to mention that these notions of proportionality are also not inherited by $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$. Making these two procedures not belong to the same class as the Phragmén-like rules. There is a superficial similarity between the intuitive explanation of the two ABCC procedures and the Phragmén-like rules. Both aim to distribute *power* to decide about the winner set in a fair way. The difference remains that Phragmén-like rules define groups based on preferences, whereas the ABCC model does not do this.

We now consider a notion of proportionality that is similar to the way Phragmén’s rule functions as it tries to assign power to equally-sized groups with a shared preference. The big difference is that it can be applied in fewer cases. The formulation is from Sánchez-Fernández et al. (2017).

Axiom 11. We say that a committee W satisfies *perfect representation* if the set of voters can be divided into k equally-sized disjoint groups $N = N_1 \cup N_2 \cup \dots \cup N_k$ ($|N_i| = \frac{n}{k}$ for each $i \in [k]$) and if we can assign a distinct candidate from W to each of these groups in a way that for each $i \in [k]$ the voters from N_i all approve their assigned candidate. An ABCC procedure $\mathcal{R}^\rho$ satisfies this property if $\mathcal{R}^\rho$ returns committees satisfying this property whenever such a committee exists.

From the rules we discussed, only Monroe’s rule, leximax-Phragmén and var-Phragmén satisfy this axiom (Brill et al., 2023; Lackner and Skowron, 2023). This notion of proportionality is incompatible with EJR and weak and strong Pareto efficiency, as can be found in Sánchez-Fernández et al. (2017). The idea is intuitively clear, as Pareto efficiency characterises rules that outputs winner sets with candidates approved by the most voters, whereas PR emphasises diversity and representation. In spirit, PR is closer to descriptive representation than EJR, as it tries to assign unique representatives to equal groups. The problem, however, is that the groups that are formed on the basis of preferences can be spread out over different communities. Put differently, each N_i that all approve of the same candidate

do not have to be contained in the same community. This observation, combined with others, leads to the following negative result:

Proposition 3.11. $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ inherit neither perfect representation nor justified representation.

Proof. Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = \{2, 2\}$. Consider the following profile:

$$\begin{array}{c} 1 \times \{a, b\} \\ 1 \times \{c, d\} \\ \hline 1 \times \{a, b\} \\ 1 \times \{c, d\} \end{array}$$

We first show that it does not inherit perfect representation. Note that it is possible to partition the voters into k equally sized groups, $N(\{a, b\})$ and $N(\{c, d\})$. If $W \in \{a, b\} \times \{c, d\}$, we can assign a distinct candidate from W to each candidate. Remark, however, that $\mathcal{R}_{RSP}^\rho$ also returns $\{a, b\}$ and $\{c, d\}$, since for $k = 1$ Monroe is the same as AV. Remark that for this example, JR and PR coincide, so it also is a counterexample for $\mathcal{R}_{RSP}^\rho$ not inheriting JR.

$\mathcal{R}_{NWP}^\rho$ would lead to the same result, since it would rerun:

$$\begin{array}{c} 2 \times X \\ 2 \times X' \end{array}$$

Where $X, X' \in \binom{C}{2}$. This would lead to the same winner sets as $\mathcal{R}_{RSP}^\rho$. □

Note that the counterexample from this proof can also be used that neither procedure inherits laminar proportionality or priceability. Both procedures inheriting none of these proportionality axioms is a bad result. Justified representation is a particularly weak notion of proportionality, that is implied by EJR. This is not satisfied. PR and EJR are incompatible, which can be seen from this example:

Example 3.4. Let $n = 8$, $m = 6$ and $k = 4$. Consider:

$$\begin{array}{ll} 1 : \{a\} & 5 : \{a, e, f\} \\ 2 : \{b\} & 6 : \{b, e, f\} \\ 3 : \{c\} & 7 : \{c, e, f\} \\ 4 : \{d\} & 8 : \{a, e, f\} \end{array}$$

A rule that satisfies PR would output $\{a, b, c, d\}$. Voter 5 – 8 from a 2-coalition, however, whereas each voter from this group only get one representative in the final winner set. Hence, this is a case where PR is but EJR is not satisfied. ◇

This example shows that the procedures do not inherit two incompatible notions of proportionality. This section was devoted to discussing the main classes of proportionality in the context of multi-winner voting. We see that the two procedures inherit no axiom from all classes. This is a negative but not a surprising result. The forced representation of some groups of voters necessarily leads to some preferences being overshadowed. Possible *escape routes* from this (and other) negative result(s) are discussed in Chapter 4.

3.5 Strategic voting

Strategic voting is studied and understood well in many social choice contexts (Conitzer and Walsh, 2016). In the context of preference aggregation, there is the seminal results by Gibbard (1973) and Satterthwaite (1975). This shows that it is impossible to have a rule that is both immune to manipulation by voters and satisfies some other basic properties. Therefore, to put it bluntly, manipulability of voting rules is something we have to live with. This is less so the case for ABC voting, as there is one natural voting rule satisfying this property: AV. This can be seen as a positive result, since AV satisfies many other desired properties — see Table 3.1. Proportionality and strategyproofness, however, are inherently impossible to satisfy simultaneously. For an in-depth discussion of the compatibility of proportionality and manipulability, the reader is referred to Peters (2018, 2019). In this section, we see that both procedures inherit strategyproofness. This means that the incompatibility of representing voters and non-manipulability is less apparent in the case of descriptive representation.

To investigate strategyproofness in the context of irresolute multi-winner voting rules, we have to define a notion of preferences over sets of objects. I base this on the work of Botan (2021). She bases these results and definitions on papers by Barberà (1977), Barberà and Pattanaik (1984), Fishburn (1972), Gärdenfors (1976), Kannai and Peleg (1984), and Kelly (1977). Since the sets contain candidates that are approved or disapproved by a certain voter, we aim to lift the preferences from candidates to sets. Hence, such a ranking is called a *preference extension* or just *extension*.

Recall that for candidates we write $c \succeq_i c'$ if $c \succ_i c'$ or $c \sim_i c'$. This captures the preference of c over c' for voter i . For sets of candidates we will write $C \overset{\circ}{\succeq}_i C'$. The following is a general and basic condition on preferences over sets of object obtained from Botan (2021).

Definition 3.4. A preference extension $\overset{\circ}{\succ}_i$ is *reflective* if $A \overset{\circ}{\succ}_i B$ implies that there is some $a \in A$ and $b \in B$ such that $a \succ b$. $\overset{\circ}{\succ}_i$ is called *strongly reflective* if $A \overset{\circ}{\succ}_i B$ implies there is some $a \in A$ and $b \in B$ such that $a \succ b$ and $\{a, b\} \not\subseteq A \cap B$.

Reflectiveness captures the idea that if a set is preferred to another, then there must exist at least one element in the preferred set that is preferred to an element in the other set. Strong reflectiveness is a strengthening of this notion. There are many well-studied notions of preference over sets of objects. Most notably *Kelly*, *Gärdenfors* and *Fishburn* preferences. Botan (2021) shows that all these notions satisfy strong reflectiveness. Therefore, if we can show that a procedure $\mathcal{R}^\rho$ inherits strategyproofness with respect to strongly reflective preferences, it also inherits strategyproofness with respect to the three aforementioned notions of preference over sets of objects. First, I show that $\mathcal{R}_{AV}$ satisfies strategyproofness with respect to strongly reflective preferences. Before this is possible, we define strategyproofness. Note that A^i is an i -variant of A if $A(j) = A^i(j)$ for all $j \in N \setminus \{i\}$ with $j \neq i$.

Axiom 12. An ABCC procedure $\mathcal{R}^\rho$ satisfies *strategyproofness with respect to strongly reflective preferences* if for all profiles A and A^i , where A^i is an i -variant of A , any community $\mathcal{Q}$ and for all $k \geq 1$ it holds that $\mathcal{R}^\rho(A, \mathcal{Q}, k) \overset{\circ}{\succeq}_i \mathcal{R}^\rho(A^i, \mathcal{Q}, k)$, where $\overset{\circ}{\succ}$ is strongly reflective.

Now, I show that this property holds for $\mathcal{R}_{AV}$.

Lemma 3.12. $\mathcal{R}_{AV}$ satisfies strategyproofness with respect to strongly reflective preferences.

Proof. Fix a voter i . Note that i prefers $\mathcal{R}_{AV}(A^i, k)$ to $\mathcal{R}_{AV}(A, k)$ if there is a winner set $W \in \mathcal{R}_{AV}(A^i, k)$ and a set $W' \in \mathcal{R}_{AV}(A, k)$ such that $W \succ_i W'$ and $\{W, W'\} \not\subseteq \mathcal{R}_{AV}(A^i, k) \cap \mathcal{R}_{AV}(A, k)$. Consider

W' . i can only add a candidate to W' , by removing a candidate she truly approves of. In this case, her satisfaction does not increase. If i approves another candidate, she also does not increase her satisfaction, since she can now only replace a candidate she truly approves of with a candidate she does not approve of or with another candidate she approves of. We see that removing a candidate from her ballot does not increase her satisfaction, neither does adding a candidate. Hence, there does not exist a W such that $W \succ_i W'$. Therefore, $\mathcal{R}_{AV}$ satisfies strategyproofness. $\square$

The following result contains extensions I have not formally defined. The corollary is mentioned, however, as it shows that $\mathcal{R}_{AV}$ satisfies strategyproofness with respect to many well-studied and appealing extensions. The definitions can be found in Chapter 2 of Botan (2021), but also in the original sources that are mentioned in the introduction to this section.

Corollary 3.13. *$\mathcal{R}_{AV}$ is strategyproofness with respect to optimistic, pessimistic, Kelly, Fishburn and Gärdenfors preferences.*

It is fairly straightforward to see that $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ inherit strategyproofness with respect to strongly reflective preferences. The intuition behind the proof is that if a voter i cannot increase her satisfaction inside of her community, she cannot increase her satisfaction in general, as the general result is an aggregate of the community elections.

Proposition 3.14. *$\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ are strategyproof with respect to strongly reflective preferences.*

Proof. In this proof I write ‘strategyproof’ instead of ‘strategyproof with respect to strongly reflective preferences’. Suppose that ρ consists of voting rules that are all strategyproof. Fix a voter i .

First, we consider $\mathcal{R}_{RSP}^\rho$. Suppose that i is part of community j , i.e., $i \in N_j$. Since $\mathcal{R}_j$ is strategyproof, i cannot change her ballot and increase her satisfaction. It is furthermore clear that she cannot influence the winner sets of any of the other communities. Since the final winner set, W , is the set of all possible (extended) unions of the community winner sets, she can only increase her satisfaction with respect to W through increasing her satisfaction with respect to W_j , but we just showed that this is impossible. Hence, $\mathcal{R}_{RSP}^\rho$ is strategyproof.

Now, we consider $\mathcal{R}_{NWP}^\rho$. As the first s rules are strategyproof, no voter can increase her satisfaction by changing her truthful ballot. Since the $s + 1$ ’th rule is also strategyproof, no change in community winner sets can lead to an increase in satisfaction for any of the voters. Hence, this procedure is also strategyproof. $\square$

Corollary 3.15. *Both $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ inherit strategyproofness with respect to Kelly, Fishburn and Gärdenfors preferences.*

Since strategyproofness is an appealing and sought-after property, its inheritance can be seen as a positive result for $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$. On the other hand, there are no natural voting rules satisfying this property. Out of the ones we discussed, only $\mathcal{R}_{AV}$ satisfies it (Lackner and Skowron, 2023). In fact, any *weakly* proportional and weakly efficient voting rule does not satisfy strategyproofness (Peters, 2018). Note that weak proportionality is implied by all the proportionality axioms we have discussed. The main idea behind this is that a voting rule that wants to represent you, can be ‘influenced’ more easily. Strategyproofness, on the other hand, requires voting rules to be ‘non-influencable’. In Chapter 4, we see that for some restricted domains it is possible to combine strategyproofness and different forms of proportionality in the ABCC model.

3.6 Computational properties

This section is devoted to understanding the computational properties of the two voting procedures. The most important computational question regarding voting rules is: *how hard is it to compute the winners for a rule (or procedure) $\mathcal{R}$?* This problem is called WINNER DETERMINATION. For ABC rules, it is formulated as follows (Aziz et al., 2015):

NAME: $\mathcal{R}$ -WINNER DETERMINATION (WD).

INPUT: An election (A, k) over a set of m candidates C .

QUESTION: What are the winning sets, $W \subseteq C$, with $|W| = k$ when applying $\mathcal{R}$ to (A, k) ?

Note that this is a *search* problem. Furthermore, note that WD for AV is polynomial-time computable. For each candidate $c \in C$, we compute $|N(c)|$. The k candidates for which this set is the largest, are returned. It can be shown that WD for a big group of Thiele methods is NP-hard (Skowron et al., 2016). Aziz et al. (2015) show that WD for PAV is NP-complete, by reducing from INDEPENDENT SET and Procaccia et al. (2008) show that CC is NP-hard. For a general overview of computational complexity theory, the reader is referred to Arora and Barak (2009).

A desirable property of the two ABCC procedures would be that they inherit the computational properties of the voting rules in ρ . In other words, we would like the class of voting rules $\mathcal{R}^\rho$ to be tractable for *any* choice of ρ . We will see that this is not the case and that computing a winner set for both procedures is intractable, even if the underlying rules are tractable. As both rules are two-step procedures and, therefore, consist of a sequence of different algorithms we analyse the computational properties of both parts separately. The first part of both rules consists of determining the community winner sets and the second part consists of some operation on these winner sets. For each part, we define a problem. This is the problem for the first part:

NAME: $\mathcal{R}^\rho$ -COMMUNITY WINNER DETERMINATION (CWD).

INPUT: An ABCC election instance $E = (A, \mathcal{Q}, k)$ over a set of m candidates C and a vector $\mathbf{q}$ of weights of length $s = |\mathcal{Q}|$.

QUESTION: For each community $N_i \in \mathcal{Q}$ where $i \in [s]$, what are the winning sets $W_i \subseteq C$, when applying $\mathcal{R}^\rho$ to $(A, \mathcal{Q}, k)$?

This problem aligns with the first step of both two-step procedures. Note that both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$ ‘inherit’ the computational properties of the underlying rules for this problem. A straightforward algorithm has the following structure:

1. Construct the community elections. $\{(A_i, k)\}_{i \in [s]}$ for $\mathcal{R}_{NWP}^\rho$ and $\{(A_i, q_i)\}_{i \in [s]}$ for $\mathcal{R}_{RSP}^\rho$.
2. For each $i \in [s]$, apply $\mathcal{R}_i$ to community elections E_i . Collect all sets $\mathcal{R}_i(E_i)$ and return them.

For any non-trivial voting rule, the second step is costliest computational task, as the first task amounts to reading the full profile A and per entry $A(i)$ adding it to a community profile A_i , which can be done in linear time. The second step does ‘inherit’ the runtime of the underlying rules, as we apply each rule in a sequential order. This means that, in the worst-case, there exists a sequential algorithm for $\mathcal{R}_{NWP}^\rho$ -CWD and $\mathcal{R}_{RSP}^\rho$ -CWD whose runtime is the sum of the runtimes of the rules in ρ . Note, again, that this is only the case if there is a rule in ρ that runs at least in linear time. Theoretically, it

is conceivable that sequentially running two hard rules is not hard, but efficient. However, it is unlikely that this is the case in general and it would be a surprising result.

For scoring rules, it is the case that we determine the score for each candidate in each community. Phragmén-like rules behave differently. Surprisingly, there exists an algorithm that is not sequential in nature but where each community election is done simultaneously if ρ consists of all $\mathcal{R}_{seqPhr}$'s. It is an open question whether the output of the simultaneous election is always the same as the sequential procedure. It will become clear that this is not a problem, as the simultaneous algorithm for $\mathcal{R}_{RSP}^\rho$ makes the second step redundant. This is not the case for $\mathcal{R}_{NWP}^\rho$. First, recall the continuous formulation of Phragmén's sequential rule (Voting rule 7). Now, consider the following algorithm for the case where $\rho = (\mathcal{R}_{seqPhr}, \dots, \mathcal{R}_{seqPhr})$.

Definition 3.5 (RSP and NWP with ρ consisting of all- $\mathcal{R}_{seqPhr}$'s.). Start with empty community committees $W_1, \dots, W_s$. Recall that each voter $i \in N$ starts with a budget of $b_i = 0$ at time $t_0 = 0$. Time increases for each community simultaneously. If there are community members with shared approved candidates and a combined budget of 1, i.e., there exists a $V \subseteq N_i$ such that $\sum_{i \in V} b_i = 1$ and $\bigcap_{i \in V} A(i) \neq \emptyset$, they can buy a candidate they all like. Reset their budgets to 0. Either, tie-breaking is done in a similar fashion to the regular seq-Phragmén rule. (*) Otherwise, consider:

Suppose that at time t , two groups V and V' in different communities have intersecting shared ballots, $\bigcap_{i \in V} A(i) \cap \bigcap_{j \in V'} A(j) \neq \emptyset$. Now, we break ties sequentially. In the first community, break ties using any tie-breaking method. Suppose we pick c_1 . In the second community, also break ties but never pick c_1 . In general, if we have picked $c_1, \dots, c_p$ in the previous p communities, break ties but never pick $c_1, \dots, c_p$ in the $p + 1$ 'th community. It can be the case that not picking $c_1, \dots, c_p$ means that there is no candidate the group can afford. Then, pick nothing and reset their budgets to 0. Continue until each community has a final committee of the desired size. In case there are less approved candidates than desired committee sizes extend them by picking candidates from C . Also do this sequentially in the same manner. So, do not break ties with the same candidates for different community.

Note that the procedure until (*) returns the same output as sequentially applying $\mathcal{R}_{seqPhr}$ to each community election. The procedure including the mechanism after (*), runs differently and the outputs are not the same as the sequential variant. It does decrease the computational cost of the second step of the two-step procedures, as there is no overlap between the final community committees. For $\mathcal{R}_{RSP}^\rho$, this means we only have to take the union of all winner sets. For $\mathcal{R}_{NWP}^\rho$, this means we run one election only. Brill et al. (2023) show that the output of seq-Phragmén can be computed in $\mathcal{O}(k^3 mn (\log n)^2)$ time. Note that their formulation and the formulation for multiple simultaneous elections I provide above, differs only on *what* we need to calculate per community, the number of calculations stays the same. Therefore, a naive estimate is that the runtime of this procedure is below $s\mathcal{O}(k^3 mn (\log n)^2)$. A proof of this fact is deferred to future work.

We now focus on the second part of both two-step procedures in general and ask ourselves: *how hard is it to compute the final winner sets, given the community winner sets as input*. We define the following decision problem to do so:

NAME: SET IN OUTPUT DETERMINATION FOR $\mathcal{R}^\rho$ ($\mathcal{R}^\rho$ -SIOD).

INPUT: A set of sets $\{\mathcal{R}_i(E_i, q_i)\}_{i \in [s]}$ containing elements from a candidate set C . A subset $S \subseteq C$.

QUESTION: Is S a subset of any of the winning sets?

Note that, technically speaking, this problem is not dependent on a voting procedure $\mathcal{R}^\rho$ but a tuple of voting rules ρ with specified weights and an operation on a collection of sets F . However, it is possible to define it for a $\mathcal{R}^\rho$ as this is defined exactly as a procedure that first determines weights, then applies ρ voting rules and does an operation on sets. The procedures RSP and NWP also restrict what the sets in the input look like. For instance, for NWP all sets are of the same size k . Lastly, if we take S such that $|S| = k$, we ask if S is a winning set. Hence, a restricted version of this problem is the decision variant of winner determination for two-step ABCC procedures.

For RSP, this problem can be reduced to the *set-covering problem*, which is NP-complete (Karp, 1972). The problem input is a *universe-set* U and a collection of sets $\mathcal{V}$. A set-cover is a union of sets from $\mathcal{V}$ such that this union is equal to U ; put formally, $\bigcup_{V \in \mathcal{V}} V = U$. For example, if $U = \{a, b, c, d, e\}$ and $\mathcal{V} = \{\{a, b, c\}, \{c, d\}, \{b, d\}, \{d, e\}\}$, then $\bigcup \mathcal{V}$ is a set-covering, but also $\{\{a, b, c\}, \{d, e\}\}$.

The decision variant of this problem is given as follows:

NAME: SET COVERING (SetCover).

INPUT: A pair $(U, \mathcal{V})$, where U is the universe-set and $\mathcal{V}$ is a collection of sets, and $p \in \mathbb{N}$.

QUESTION: Does there exist a subset $V \subseteq \mathcal{V}$ such that $|V| = p$ and $\bigcup V = U$?

We now consider RSP-SIOD. The core idea of the following proof is that we reduce the collection of sets $\mathcal{V}$ from the input of SetCover to sets of sets $\mathcal{R}_i(E_i, q_i)$. For this reduction to work, it must be the case that all elements of any $\mathcal{V}$ could be the output of a voting rule in ρ . If this is not the case, we do not reduce to an instance of $\mathcal{R}^\rho$ -SIOD. Theoretically speaking, however, this is trivial, as the voting rules in ρ can be any function mapping into $\wp(C)$. The interesting question, however, is whether there is a natural voting rule in which the hardness kicks in. I show by an example that this indeed is the case. First, the general result is presented.

Proposition 3.16. *RSP-SIOD is NP-hard.*

Proof. Hardness is shown by reducing from SetCover. The input is a set U , a collection of sets $\mathcal{V}$ and an integer k . Note that we have to reduce this to an instance with a candidate set C and a collection of s output sets $\{\mathcal{R}_i(E_i, q_i)\}_{i \in [s]}$, for brevity write r_i instead of $\mathcal{R}_i(E_i, q_i)$. Define, $s := p$.

Now, let l be the cardinality of the largest set in $\mathcal{V}$. Extend $\mathcal{V}$ to $\mathcal{V}'$ by adding $l - |V|$ ‘filler candidates’ $x_1, \dots, x_{l-|V|}$ for each $V \in \mathcal{V}$. This way, all sets in $\mathcal{V}'$ are of the same size, which is a requirement on all sets from r_i for each $i \in [s]$. For each $i \in [p]$ we define $r_i := \mathcal{V}'$.

C can be constructed by adding all elements from $\bigcup \mathcal{V}$ and the filler candidates.

Lastly, we define $S := U$. S is the set of which we want to know if this is a subset of any of the final winner sets. Note that these are steps that can all be done in polynomial time. Now, if S is an element of any of the final winner sets, there exists a set-covering of size p . $\square$

This reduction can be clarified by applying it to an example.

Example 3.5. Let the input of the SetCover instance be $U = \{a, b, c, d, e\}$, $\mathcal{V} = \{\{a, b, c\}, \{d, e\}, \{c, d\}, \{b, d\}\}$ and $p = 2$. There are $s := p = 2$ communities. $\mathcal{V}'$ can be constructed by adding $|\{a, b, c\}| - |V|$ elements to each $V \in \mathcal{V}$. Hence, we obtain $\mathcal{V}' = \{\{a, b, c\}, \{d, e, x_1\}, \{c, d, x_1\}, \{b, d, x_1\}\}$. The output of the two communities is given as follows:

$$\frac{\{\{a, b, c\}, \{d, e, x_1\}, \{c, d, x_1\}, \{b, d, x_1\}\}}{\{\{a, b, c\}, \{d, e, x_1\}, \{c, d, x_1\}, \{b, d, x_1\}\}}$$

$S = U = \{a, b, c, d, e\}$ and $C = \{a, b, c, d, e, x_1\}$. One of the final winner sets is $W = \{a, b, c, \} \cup \{d, e, x_1\}$, which contains S . This winner set is the union of the k sized set-cover.

It is possible to show that the aforementioned winner sets are the outputs of existing voting rules for a given election. Let $n = 9$, $m = 5$ and $k = 3$. Consider the following profile:

1 : $\{a, d\}$	4 : $\{b, d\}$	7 : $\{c, d\}$
2 : $\{a, b, c, e\}$	5 : $\{b, c, e\}$	8 : $\{b, c, e\}$
3 : $\{a, x_1\}$	6 : $\{b, x_1\}$	9 : $\{c, x_1\}$

Take $\mathcal{R}_{CC}$ or seq-CC as the voting rule. The four sets in $\{\{a, b, c\}, \{d, e, x_1\}, \{c, d, x_1\}, \{b, d, x_1\}\}$ are the only ways in which it is possible to represent every single voter. $\diamond$

Note that this is not a general recipe for turning any set of sets from the reduction into the output of an election instance to which we apply a ‘natural’ voting rule. To my knowledge, it is an open question whether such a general recipe exists. As Faliszewski and Skowron (2017) show that finding the output of $\mathcal{R}_{CC}$ is the same as finding a solution to the *maximum coverage problem*, which is closely related to SetCover, using $\mathcal{R}_{CC}$ as the natural rule to find such a general recipe seems to be a good starting point.

As a contextualisation: note that hardness of RSP-SIOD is a strong indication for $\mathcal{R}_{RSP}^\rho$ being intractable, but it is not a full proof of this fact. For such a proof, one would have to define a related problem in which the input is altered. Instead of a set of outputs of voting rules $(\{\mathcal{R}_i(E_i, q_i)\}_{i \in [s]})$, community election instances would have to be added and a tuple of rules ρ would be part of the context. The reason behind this is that there might exist a tuple of tractable rules for which it is not possible to construct the set of sets needed for the reduction from SetCover. The example above, however, is an indication for there being a tuple of efficient rules for which RSP is intractable, in the worst case.

Note that NWP-SIOD reduces to multiple instances of winner determination for $\mathcal{R}_{s+1}$ on a restricted class of profiles. That is, all the profiles $\Pi \in \mathbf{A}$. For some choices of $\mathcal{R}_{s+1}$ and elections $E = (A, \mathcal{Q}, k)$ this problem is tractable. If we apply $\mathcal{R}_{AV}$ at the last step and there is one community $N_i \in \mathcal{Q}$ that is the majority, their approval ballot is always the final output and this can be determined by inspecting their community winner sets. As the second step for NWP is dependent upon a voting rule $\mathcal{R}_{s+1}$ I do not provide a general result. Moreover, it is likely that it is intractable as we can define a highly intractable rule and use it as $\mathcal{R}_{s+1}$. Constructing the profiles, furthermore, is also a computationally expensive operation as the most natural operation to do so is taking the Cartesian product. Given these facts, it seems unlikely that there is a general and efficient algorithm that solves NWP-SIOD. Therefore, I omit a further investigation.

Lastly, note that the given results are *worst-case results*, meaning that there are domains for which the decision problem has an efficient algorithm. Consider, the case where C can be partitioned into s sets, $C_1, \dots, C_s$ and each community N_i approves only of candidates in $C_i \subset C$ and does not approve of any $C \setminus C_i$. For RSP, this amounts to checking whether S is a subset of any of the community winner sets.

3.7 Discussion

Table 3.1 contains an overview of all axioms that are inherited by each procedure. The last column shows the computational complexity of $\mathcal{R}^\rho$ -SIOD for RSP and NWP. For the ABC rules, we show

winner determination. The table also shows which properties are *satisfied* for some of the most important voting rules we discussed. This is based on a table from the survey of Lackner and Skowron (2023).

	$\mathcal{R}_{RSP}^\rho$	$\mathcal{R}_{NWP}^\rho$	$\mathcal{R}_{AV}$	$\mathcal{R}_{CC}$	$\mathcal{R}_{PAV}$	$\mathcal{R}_{Monroe}$	$\mathcal{R}_{seqPhr}$
Weak efficiency	No ⁽¹⁾	No ⁽¹⁾	Yes	Yes	Yes	Yes	Yes
Pareto efficiency	No	No	Strong	Weak	Strong	No	No
Support monot. with add. voters	?	?	Yes	Yes	Yes	No	Cand. ⁽²⁾
Support monot. without add. voters	Yes	Yes	Yes	Cand. ⁽²⁾	Cand. ⁽²⁾	Cand. ⁽²⁾	Cand. ⁽²⁾
EJR	No	No	No	JR	Yes	JR	Yes
Perfect JR	No	No	No	No	No	Yes	No
Committee monotonicity	Yes	No	Yes	No	Yes	No	Yes
Strategy- proofness	Yes	Yes	Yes	?	No	No	No
External consistency	No	No	Yes ⁽³⁾	Yes ⁽³⁾	Yes ⁽³⁾	No ⁽³⁾	No ⁽³⁾
Complexity	NP ⁽⁴⁾	? ⁽⁴⁾	P	NP	NP	NP	P

Table 3.1: (1): This holds only for the original formulation. (2): Candidate monotonicity is defined as support monotonicity for candidates, not sets of candidates. So, $X = \{c\}$ where $c \in C$. (3): For ABC rules, this should be read as ‘consistency’. (4) Recall that we can only justifiably believe that NWP and RSP are intractable.

Interesting to note about this table is that Pareto efficiency and PR are incompatible and both are not satisfied. Pareto efficiency can be seen as an axiom that prefers winner sets that are preferred by larger groups of voters, whereas the proportionality axioms, like EJR, give up excellence of candidates for diversity of voter representation. If the ABCC model has some representative properties, these are not characterised by the proportionality axioms we have seen in this chapter. Since RSP can be seen as running multiple rules for parts of the profile and adding them together, all monotonicity axioms and strategyproofness are inherited. The short reason being that these axioms are not influenced by what happens in other communities. In some sense, these are *community-local* properties with respect to RSP. I expect, however, that we do not keep these axioms if we want to construct a procedure that is made resolute by breaking ties.

The final verdict for NWP is that it is an underperforming rule and not fit for any serious applications. In Section 2.4 we saw that it gave non-proportional results. Recall Example 2.10, with $n = 2$, $m = 4$, $k = 2$ and $\mathcal{Q} = [1, 2]$:

$$\frac{1 \times \{a, b\}}{1 \times \{a, b\} + 1 \times \{c, d\}}$$

For both $\rho = (\mathcal{R}_{AV}, \mathcal{R}_{AV}, \mathcal{R}_{PAV})$ and $\rho = (\mathcal{R}_{AV}, \mathcal{R}_{AV}, \mathcal{R}_{AV})$ it is the case that one of the outputs is $\{c, d\}$. This means that there is an output that is neither diverse nor proportional nor is it Pareto efficient. Combined with the fact that NWP inherits properties that are satisfied simultaneously only by $\mathcal{R}_{AV}$, this rule performs too poorly to be analysed any further.

Chapter 4

Escape routes

As we have seen in the previous chapters, both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$ perform poorly with respect to important and basic requirements. In particular in terms of proportionality, as they do not inherit any of the proportionality axioms we discussed. In their original formulation, they also do not satisfy weak efficiency. Fortunately, this was resolved with a minor modification. Secondly, determining their winner sets is computationally hard. Recall, however, that we discussed some choices of rules for which this is not the case. Lastly, they give counter-intuitive outputs in terms of (descriptive) representation, as we saw in Section 2.4. This chapter contains a discussion of possible *escape routes* for these problems. Note that they are not full ‘solutions’, since providing a full axiomatic and computational analysis of the possible solutions goes beyond the scope of this thesis. Hence, what we will encounter are ‘escape routes’.

The first escape route is based on the idea that we should not *always* use the ABCC model. Only when there is a *meaningful* way to organise the set of voters into communities. This will lead to a discussion of some basic *domain restrictions*. The second escape route is the formulation of new voting rules that, ideally, give intuitive verdicts that balance descriptive and substantive representation for the examples in Section 2.4 and whose winner sets can be computed efficiently. We also consider new axioms specific to the ABCC model. This is based on the idea that the proportionality axioms we have considered so far are not suited for the ABCC situation. Towards the end, I discuss which of these escape routes is the most promising. Note that the three escape routes do not appear sequentially. The new axioms will be defined based on observations we make while defining the domain restrictions and voting rules. But first, we (re)consider the possibility that the ABCC model is suited for neither descriptive representation nor substantive forms of proportionality, such as those guaranteed by EJR and PR. This is the most cynical and fourth solution: abandoning the ABCC model.

We start by looking at an example, which I will refer to by the *fatal flaw* of both rules. As mentioned in the discussion towards the end of chapter 3, $\mathcal{R}_{NWP}^\rho$ will not be discussed, only $\mathcal{R}_{RSP}^\rho$.

Example 4.1 (Fatal flaw). Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = [2, 2]$. Consider:

$$\begin{array}{l} 1 : \{a, b\} \\ 2 : \{c, d\} \\ \hline 3 : \{a, b\} \\ 4 : \{c, d\} \end{array}$$

Recall that both $\mathcal{R}_{NWP}^\rho$ and $\mathcal{R}_{RSP}^\rho$ have the same output as $\mathcal{R}_{AV}$ here, i.e., $\binom{C}{2}$. Recall also that $\mathcal{R}_{AV}$ is seen as one of the least proportional scoring rules. $\diamond$

We have already established that all proportionality axioms we have encountered, EJR and PR, pick a set from $\{a, b\} \times \{c, d\}$. $\mathcal{R}_{RSP}^\rho$ does not do this, as it also outputs $\{a, b\}$ and $\{c, d\}$. This can mean two things. Either, the procedure captures a form of representation EJR and PR do not, or RSP does not do this and it fails in terms of proportionality. My claim is that the second is the case. I argue from the contrary. Suppose that $\mathcal{R}_{RSP}^\rho$ picking $\{a, b\}$ satisfies a form of descriptive representation in the case of the fatal flaw. I only discuss $\{a, b\}$, as it is analogous to discussing $\{c, d\}$. First, I claim that the following situations are the *only* possible ways in which $\mathcal{R}_{RSP}^\rho$ can provide a form of descriptive representation that EJR and PR do not enforce:

Example 4.2. Let $n = 4$, $m = 4$, $k = 2$ and $\mathcal{Q} = [2, 2]$. Consider:

$(1, \alpha) : \{(a, \alpha), (b, \beta)\}$	$(1, \alpha) : \{(a, \alpha), (b, \beta)\}$	$(1, \alpha) : \{(a, \alpha), (b, \beta)\}$
$(2, \alpha) : \{(c, \alpha), (d, \alpha)\}$	$(2, \alpha) : \{(c, \beta), (d, \alpha)\}$	$(2, \alpha) : \{(c, \alpha), (d, \beta)\}$
$(3, \beta) : \{(a, \alpha), (b, \beta)\}$	$(3, \beta) : \{(a, \alpha), (b, \beta)\}$	$(3, \beta) : \{(a, \alpha), (b, \beta)\}$
$(4, \beta) : \{(c, \alpha), (d, \alpha)\}$	$(4, \beta) : \{(c, \beta), (d, \alpha)\}$	$(4, \beta) : \{(c, \alpha), (d, \beta)\}$

There are three profiles in total. This model is not formally defined, but the tuples denote the voter or candidate and the group, which are denoted by α and β , they belong to. This is similar to the AAV model, but with attributes of voters added. Voters belonging to the same community share the same attribute. In the case of descriptive representation, it makes sense to only consider the case where the candidates belong to the same groups as the voters. So, all candidates and voter belong to α or β .

This example is exhaustive, since $W = \{a, b\}$ provides a form of descriptive representation if a and b belong to different communities. Without loss of generality, we assume that a belongs to α and b to β . The example above contains all possible ways of dividing α and β to the candidates such that a and b belong to different communities. Symmetric cases, where c and d both belong to β , for instance, or the α and β 's are swapped, have been left out. In all these examples, $\{a, b\}$ leads to descriptive representation, but there is always an argument to be made for choosing $\{b, c\}$ or $\{b, d\}$, as these also provide descriptive representation and more voters have a candidate of their liking in the final winner set. Hence, in the case of the fatal flaw, $\mathcal{R}_{RSP}^\rho$ does not provide a form of descriptive representation that EJR and PR do not. $\diamond$

In general, the fatal flaw shows us that $\mathcal{R}_{RSP}^\rho$ performs poorly if there is no correlation between voting behaviour and group membership. When the profile of each community is the same, descriptive representation should be put aside for a substantive form of representation, such as EJR or PR. The reason behind this is that choosing a candidate from a particular community, while ignoring the preferences of members of that community, does not lead to that community having the *power* to decide what the winner set is. Recall that in Section 2.1 we set out that *empowering* communities is one of the goals of the ABCC model. This observation leads to the following ABCC property:

Axiom 13 (Scaling property (SP)). Let $E = (A, \mathcal{Q}, k)$ be an election instance. If all communities $Q \in \mathcal{Q}$ have the same ballot $A|_Q$, then $\mathcal{R}^\rho(A, \mathcal{Q}, k) \subseteq \mathcal{R}^\rho(A, [n], k)$ for any $\mathcal{Q} \in \mathcal{Q}$.

This captures the case where we ‘rescale’ the larger election into s smaller elections. The fatal flaw is a counterexample to $\mathcal{R}_{RSP}^\rho$ satisfying this property. Note that $\{a, b\} \in \mathcal{R}^\rho(A, \mathcal{Q}, k)$ if ρ consists of $\mathcal{R}_{PAV}$'s only. But $\mathcal{R}^\rho(A, [n], k)$ contains only sets constructed from $\{a, b\} \times \{c, d\}$. Note that, in this chapter, we have only considered the fatal flaw to argue in favour of the ABCC model. The fatal flaw

is an extreme case, in which there is nothing distinguishing the communities and even here the ABCC model points towards something interesting: if demographic traits and preferences are completely independent, we should pursue substantive representation. This idea will be formalised in the following chapter.

4.1 Conditions and (domain) restrictions

This section is devoted to the first escape route: formulating restrictions and conditions such that $\mathcal{R}_{RSP}^\rho$ does inherit EJR or PR. First, the case where each community has their *own unique representatives* is analysed. Second, the focus is shifted towards the case where voters belonging to different communities may share candidates. Note that voters in a group V *share* a candidate c if $c \in \bigcap_{i \in V} A(i)$. There are three kinds of results in this section. Recall that one of the motivations of the ABCC model was to investigate whether it is possible to structure ballots in such a way that applying a non-proportional rule in each community, leads to a proportional output. The first class of results is related to this motivation. I investigate the conditions on the community structure which lead to a proportional result when ρ consists of $\mathcal{R}_{AV}$'s only. The second kind of result are conditions and restrictions that lead to inheritance or satisfaction of EJR. Third, the same for perfect representation. I first define what it means for a community structure $\mathcal{Q}$ to be *disjoint*.

Definition 4.1. Let $E = (A, \mathcal{Q}, k)$ be an election. The community structure $\mathcal{Q}$ is *disjoint* and every community has their own distinct representatives, i.e., for any $Q, Q' \in \mathcal{Q}$ it holds that

$$\bigcup_{i \in Q} A(i) \cap \bigcup_{j \in Q'} A(j) = \emptyset$$

Even if all communities are disjoint, EJR is not inherited. We saw this in chapter 3 already, the reason behind this is that there might be an instance of ‘over-representation’. It is possible to construct a general counterexample based on this intuition as follows:

Example 4.3. Fix $\ell \in \mathbb{N}$ as an even number. Pick $n = 2 \cdot \ell + \frac{\ell}{2}(s - 1)$, $k = s$ and $m = s + 1$ where s is uneven. Now, construct:

$$\begin{array}{c} 2 \cdot \ell \times X_1 \\ \hline \frac{\ell}{2} \times X_2 \\ \hline \vdots \\ \hline \frac{\ell}{2} \times X_s \end{array}$$

Hence, $|\mathcal{Q}| = s$. X_1 contains c_1 and c_2 . For $i > 1$, define $X_i := \{c_{i+1}\}$. So, $X_i \cap X_j = \emptyset$ if $i \neq j$. The ‘upper’ community contains a 2-cohesive group but $q_i = 1$ for all $i \in [s]$, by diversity of the weights. Hence each member of the 2-cohesive group only gets one candidate in the final winner set, which does not satisfy EJR. Applying $\mathcal{R}_{PAV}$ on all communities would lead to the same output as $\mathcal{R}_{AV}$. So, this is also a counterexample to inheritance of EJR for disjoint communities if $|\mathcal{Q}| < \frac{n}{k}$. $\diamond$

The reason why $\mathcal{R}_{RSP}^\rho$ fails EJR for examples like these is similar to why $\mathcal{R}_{CC}$ fails EJR. $\mathcal{R}_{CC}$ is a *diverse* rule, whereas $\mathcal{R}_{PAV}$ is a *proportional* rule. Which means that $\mathcal{R}_{PAV}$ grants candidates to *sufficiently large* groups, whereas $\mathcal{R}_{CC}$ tries to maximise the number of people represented; ‘larger’ groups get as many representatives as ‘smaller’ groups. To highlight this, consider this example:

Example 4.4. Fix $\ell = 2$ and pick $s = 3$. So, $n = 6$, $m = 4$ and $k = 3$.

$$\begin{array}{c} 2 \times \{a, b\} \\ \hline 1 \times \{c\} \\ \hline 1 \times \{d\} \end{array}$$

The output is $\{\{a, c, d\}, \{b, c, d\}\}$, whereas only $\{a, b, c\}$ or $\{a, b, d\}$ satisfy EJR. The general output is diverse, but it is not proportional. Hence, in terms of representation the output is not bad. $\diamond$

This observation leads to the following result. Note that we do not assume disjointness of $\mathcal{Q}$ for this.

Theorem 4.1. *Let $E = (A, \mathcal{Q}, k)$ be an election. For all communities, if all voters share one candidate, the output of $\mathcal{R}_{RSP}^\rho$, with ρ consisting of all- $\mathcal{R}_{AV}$'s, is contained in $\mathcal{R}_{CC}(A, k)$.*

Proof. Note that the profile has the following structure:

$$\begin{array}{c} 1 \times \{c_1\} \cup X_1^1 \\ \vdots \\ 1 \times \{c_1\} \cup X_{n_i}^1 \\ \hline \vdots \\ 1 \times \{c_s\} \cup X_1^s \\ \vdots \\ 1 \times \{c_s\} \cup X_{n_s}^s \end{array}$$

Here X_j^i is the set of approved candidates of voter j in community i . $\{c_1, \dots, c_s\} \cup X$, where X is any subset of $C \setminus \{c_1, \dots, c_s\}$ is included in $\mathcal{R}_{CC}(A, k)$, because every voter is represented. $\mathcal{R}_{RSP}^\rho$ with ρ containing all $\mathcal{R}_{AV}$'s does the same, but then divided into two steps. First, in each community, the shared candidate is elected, as she has the highest $\mathcal{R}_{AV}$ -score. Then, the remaining candidates with the most votes are elected in each community where the weights are greater than 1. This leads to the same set where X contains the candidates with the highest $\mathcal{R}_{AV}$ -scores from all candidates in $C \setminus \{c_1, \dots, c_s\}$. Which, as mentioned, is also contained in the output of $\mathcal{R}_{CC}$. $\square$

The assumption that each community shares one candidate is rather weak, whereas the resulting winner set is the same as for CC, which is a typical diverse rule and satisfies JR. Furthermore, recall that $\mathcal{R}_{AV}$ is strategyproof. This shows that under this mild condition, it is possible to balance a form of representation and strategyproofness. Recall from Section 3.5 that these properties are incompatible in the general case. The statement above can, therefore, be seen as a positive result.

We now consider the computational problem of recognising whether a given profile A can be divided into communities where each community shares at least one candidate. It is likely that it is NP-hard, as can be made convincing by considering the decision problem that encodes the recognition problem. Note that a collection of sets $\mathcal{V}$ is α -disjoint if for each $V \in \mathcal{V}$ it holds that $|\bigcap V| \geq \alpha$.

NAME: 1-DISJOINT RECOGNITION (1-Disjoint).

INPUT: A profile A and an integer k .

QUESTION: Does there exist a collection $\mathcal{V} = \{V_1, \dots, V_m\}$ such that $V_i = \{A(i_1), \dots, A(i_j)\}$ where $A(i_j)$ are ballots, $\bigcap_{S \in V_i} S \neq \emptyset$, $|\mathcal{V}| \leq k$ and $\bigcup \mathcal{V} = S$.

As A can be seen as a set and $\mathcal{V}$ can be seen as a covering of A with extra constraints, it is not ridiculous to believe that this problem is NP-hard, as SetCover for a fixed k is as well. The minimisation variant of SetCover, is also NP-hard. The latter is more similar to 1-Disjoint.

We now proceed with defining sufficient conditions for the satisfaction or inheritance of (E)JR. Note that for Theorem 4.1 we assumed that all candidates in a community share at least one candidate. If we change this assumption to disjointness, we obtain the following (positive) result.

Proposition 4.2. *Let $E = (A, \mathcal{Q}, k)$ be an election. Suppose that all communities are disjoint for all Q and each community shares at least 1 candidate, then:*

- *JR is satisfied whenever ρ consists of all $\mathcal{R}_{AV}$'s.*
- *JR is inherited.*

Proof. Satisfaction follows from Theorem 4.1, combined with the fact that $\mathcal{R}_{CC}$ satisfies JR.

Inheritance is also straightforward. If there is a 1-cohesive group V , it is part of one community; $V \subseteq Q$. Note that Q itself is a 1-cohesive community with respect to Q , but also with respect to N . As $V \cap Q$ is non-empty, there is at least one candidate in V that gets represented by one candidate. By the working of RSP, this candidate makes it into the final winner set. As we assumed V to be any 1-cohesive group, the all 1-cohesive groups have at least one member that gets represented. $\square$

This is a positive result because the assumptions are quite modest, and JR is a sought-after property in the context of proportionality. We see that the conditions become more tedious if we wish the generalise from JR to EJR. For instance, recall that example 4.3 showed us that we have to assume that $|\mathcal{Q}| \geq \frac{n}{k}$.

Proposition 4.3. *Let $E = (A, \mathcal{Q}, k)$ be an election. Suppose that all communities are disjoint and whenever $|Q| = \frac{n}{t}$ for any $t > 0$, then the weight of this community is $\leq \frac{k}{t}$, i.e., $q_Q \leq \frac{k}{t}$, then $\mathcal{R}_{RSP}^\rho$ inherits EJR.*

Proof. Suppose that V is a ℓ -cohesive group. This means that $|V| \geq \ell \cdot \frac{n}{k}$ and they share ℓ candidates. By disjointness it is the case that $V \subseteq Q$ for a community Q . This means that $|Q| \geq |V|$. Note that $|Q|$ can always be written as $\frac{n}{t}$ for a $t > 0$, because $|Q|$ is a rational number. Now, we know that $q_Q \leq \frac{k}{t}$. Consider:

$$|V| \geq \ell \frac{n}{k} = \ell \frac{\frac{n}{t}}{\frac{k}{t}} \geq \ell \frac{\frac{n}{t}}{q_Q} = \ell \frac{|Q|}{q_Q}$$

This shows that V is a ℓ -cohesive community with respect to the community election in Q . Hence, any rule satisfying EJR in this community elects ℓ candidates representing at least one voter in V . As these candidates reappear in all community winner sets, they also appear in the final winner sets. $\square$

$\mathcal{R}_{NWP}^\rho$ does not inherit JR for the same condition. This can be seen by observing the following counterexample:

Example 4.5. Let $n = 6$, $k = 3$, $m \geq 5$ and $\mathcal{Q} = [2, 2, 2]$. Let ρ consist of all $\mathcal{R}_{PAV}$'s. Consider:

$$\begin{array}{c} 2 \times \{a\} \\ \hline 2 \times \{b\} \\ \hline 2 \times \{c\} \end{array}$$

One of the elections $\mathcal{R}_{NWP}^\rho$ reruns is:

$$\begin{array}{l} 2 \times \{a, d, e\} \\ 2 \times \{b, d, e\} \\ 2 \times \{c, d, e\} \end{array}$$

$\mathcal{R}_{PAV}$ would output $W = \{\{a, d, e\}, \{b, d, e\}, \{c, d, e\}\}$ which does not satisfy JR. Note that it also does not satisfy weak efficiency. $\diamond$

It is a relatively basic desideratum for any procedure to inherit JR if all community ballots are disjoint and each community is of size $\geq \frac{n}{k}$. $\mathcal{R}_{NWP}^\rho$ not satisfying these properties is another argument against this procedure.

The two Propositions 4.2 and 4.3 can be seen as instances of a more general condition. To see this, one has to note that EJR is guaranteed to be inherited if a ‘global’ ℓ -cohesive group — with respect to the ABC election (A, k) — also is a ‘local’ community ℓ -cohesive group. If this is the case, a community rule satisfying EJR represents this ℓ -cohesive group with respect to some community election (A_Q, q_Q) . As RSP returns unions of community winner sets, EJR is satisfied globally. Assuming disjointness made it easier to formulate numerical conditions on the size of $|Q|$ and the weights q_Q to guarantee that a ‘global’ ℓ -cohesive group gets scaled in such a way that it is also an ℓ -cohesive group with respect to a community. To define a sufficiency-condition without assuming disjointness, we have to first define a new notion of cohesiveness:

Definition 4.2 (Internal ℓ -cohesiveness). Let $E = (A, \mathcal{Q}, k)$ be an election. Take any community $Q \in \mathcal{Q}$. A group $V \subseteq Q$ is *internally ℓ -cohesive* if:

1. $|V| \geq \ell \cdot \frac{|Q|}{k}$.
2. $|\cap_{i \in V} A_{\upharpoonright Q}(i)| \geq \ell$

Note that $A_{\upharpoonright Q}$ refers to the ballot of community Q .

It is possible to rank all such internal cohesive groups based on their size ($|V_1| \geq |V_2| \geq \dots \geq |V_j|$). This leads to the notion of a *maximally internal ℓ -cohesive group* and a *minimally internal ℓ -cohesive group*. For readability, I will call regular ℓ -cohesive groups ‘external’ ℓ -cohesive groups.

The *maximally* and *minimally* external ℓ -cohesive groups are defined analogously to the internal case. In Section 2.4 we saw some examples in which the ‘internal cohesiveness’ was smaller than the ‘external cohesiveness’, this lead to $\mathcal{R}_{RSP}^\rho$ not satisfying EJR. The ‘fatal flaw’ is also an example of this.

Example 4.6. Let $n = 3$, $m = 5$, $k = 3$ and $\mathcal{Q} = [1, 2]$.

$$\begin{array}{l} 1 : \{a, b, c\} \\ 2 : \{a, b, d\} \\ 3 : \{c, d, e\} \end{array}$$

NWP outputs $\{c, d, e\}$ with a probability of $\frac{1}{6}$. For RSP with $\mathbf{q} = (1, 2)$, the output is $\{\{a\}, \{b\}, \{c\}\} \times \{\{a, d\}, \{b, d\}, \{c, d\}, \{d, e\}\}$ which also includes $\{c, d, e\}$. This does not satisfy EJR, because voters 1 and 2 each deserve two representatives and now they only get one. $\diamond$

One way to reformulate this is that the *minimal external cohesiveness is strictly smaller than the maximal internal cohesiveness*. Voters 1 and 2 belong to different communities, but they have more in common than with their community members. Another way to interpret this is that *there exists an internal cohesive group that is not part of an external cohesive group*. In the remainder, we only consider the latter interpretation of the example and the former is relegated to future work. The latter interpretation leads to the following result:

Proposition 4.4. *Let $E = (A, \mathcal{Q}, k)$ be an election. If for any external ℓ -cohesive group V^+ there exists an internal cohesive ℓ' -group V^- such that $V^- \subseteq V^+$, JR is inherited.*

This is not a sufficient condition for EJR. As can be seen from the following counterexample:

Example 4.7. Let $n = 8$, $m = 17$, $k = 6$ and $\mathcal{Q} = [3, 3, 2]$. Consider:

$2 \times \{a, b\}$	$1 \times \{a, b\}$	$1 \times \{c_{10}, c_{11}\}$
$1 \times \{c_3, c_4\}$	$1 \times \{c_7, c_8\}$	$1 \times \{c_{12}, c_{13}\}$
	$1 \times \{c_5, c_6\}$	$1 \times \{c_{14}, c_{15}\}$

$N(\{a, b\})$ is a set of three voters. $3 > 2 \cdot \frac{8}{6}$, which means both $\{a, b\}$ should be always selected if we want to satisfy EJR. The two voters that approve of $\{a, b\}$ in the first community, are a 1-cohesive group, since $2 > \frac{3}{2}$, but not a 2-cohesive group. This means that a potential winner set is $\{a, c_3\} \cup \{c_7, c_8\} \cup \{c_{10}, c_{11}\}$, which does not satisfy EJR, but it does satisfy JR. Since the winner set of the first community does always include a or b , JR is always satisfied, in this case. $\diamond$

The proof of the statement above goes as follows. It is fairly straightforward:

Proof of theorem 4.4. If there is an internal 1-cohesive group that is part of an external 1-cohesive group, any rule satisfying JR includes a representative candidate in the community winner set. After combining, this candidate is part of all winner sets. $\square$

The results of the inheritance of EJR and the satisfaction of (E)JR when ρ consists of all- $\mathcal{R}_{AV}$'s is omitted. The reason being that after observing the systematically generated profiles, it is likely that the sufficient conditions are all instances of this intuition: *EJR is inherited if any external ℓ -cohesive group is an internal ℓ -cohesive group*, which is trivial. This, however, is not a necessary condition, as it is possible that three 1-cohesive groups each get a candidate in three separate communities that leads to a final winner set satisfying EJR for a 3-cohesive group. Satisfaction of EJR if ρ consists of all- $\mathcal{R}_{AV}$'s is guaranteed if members of an external ℓ -cohesive group get a candidate of their liking selected by $\mathcal{R}_{AV}$, the conditions for this would lead out to spelling out the workings of the $\mathcal{R}_{AV}$ -rule for different combinations of communities.

We now formulate a sufficiency-condition for $\mathcal{R}_{RSP}^\rho$ satisfying EJR if ρ consists of all- $\mathcal{R}_{AV}$'s. The condition is demanding, as the intuition behind the following results is that EJR is satisfied if every community corresponds to a ℓ -cohesive group.

Theorem 4.5. *Let $E = (A, \mathcal{Q}, k)$ be an election. Suppose that $\mathcal{Q}$ is disjoint. EJR is satisfied by $\mathcal{R}_{RSP}^\rho$ with ρ consisting of all- $\mathcal{R}_{AV}$'s if the following conditions hold:*

- $|Q| \geq \frac{n}{k}$ for all $Q \in \mathcal{Q}$
- Any $\geq \ell \cdot \frac{n}{k}$ -sized community gets ℓ seats.

- Each community $Q \in \mathcal{Q}$ shares as many candidates as they have weights.

Proof. If there is a ℓ -cohesive group it is contained in a community $Q \in \mathcal{Q}$, because of disjointness. Because of the second assumption, any ℓ -sized community gets ℓ seats. The third assumption tells us that they share ℓ candidates. These have the highest $\mathcal{R}_{AV}$ -score, so at least ℓ representatives get elected into all winner sets of Q , W_Q . As $W_Q \subseteq W$, EJR is satisfied. $\square$

From Example 4.3 it follows that if the first two conditions do not hold, EJR is not satisfied. The same can be said for the third condition, because of the fatal flaw. Recall that $n = 4 = m$, $k = 2$ and $\mathcal{Q} = [2, 2]$. Consider:

1 : {a, b}
2 : {c, d}
3 : {a, b}
4 : {c, d}

The first and second condition is satisfied, but each share 0 (which is smaller than q_i for both i) candidates. The output here does not satisfy EJR. This points towards necessity of these conditions, but a full systematic proof of this fact has to be provided.

We now reflect on the domain restrictions for inheritance and satisfaction of EJR. The results for disjoint communities are relatively positive. Disjointness is an extreme case, in which belonging to a community strongly determines the voting behaviour of its members. In these situations, $\mathcal{R}_{RSP}^\rho$ seems to work as intended. Satisfying or inheriting EJR, even if the communities are disjoint, comes with some tedious conditions, suggesting that RSP does not naturally provide the kind of proportionality EJR guarantees.

Lastly, note RSP-SIOD is tractable if $\mathcal{Q}$ is disjoint. First note that we never need to extend after combining community winner sets if communities are disjoint, as this is only necessary if two communities have overlapping winning sets. A possible algorithm for checking if a set S is part of the final winner set has the following structure:

1. Because of disjointness, S can be divided into candidates that ‘belong’ to a certain community. This can be done by, for instance, taking the union over the ballots of each community, i.e., $\bigcup_{i \in Q} A(i)$ for all $Q \in \mathcal{Q}$. After we know which candidates belong to which community, divide $S = S_1 \cup \dots \cup S_s \cup X$ according to this. If X is non-empty, we know that S contains candidates that represent no community. So, we know that S is not a subset of any of the winning sets.
2. For each $i \in [s]$, check if $S_i \subseteq W_i$ for a $W_i \in \mathcal{R}_i(A_i, q_i)$. If this is the case for each $i \in [s]$, the answer is ‘yes’. Otherwise, return ‘no’.

Note that the operations in the first step can all be done efficiently. This holds for taking an union over a set, constructing $S = S_1 \cup \dots \cup S_s \cup X$ and checking if X is non-empty. The same goes for the second step, as we only have to take s inclusions.

To finish this section, we investigate domain restrictions for perfect representation. Formulating conditions to satisfy or inherit PR is tedious. It is important to note that we do not assume any more that all communities in $\mathcal{Q}$ are disjoint. Instead, we assume a form of ‘local disjointness’. Also note that the definition of PR forces us to assume that there is a partition of N into k parts $N_1 \uplus \dots \uplus N_k$. I refer to each N_i here by ‘group’ to prevent ambiguity and for communities I will use $Q_i \in \mathcal{Q}$. Instead

of assuming that all communities $Q, Q' \in \mathcal{Q}$ are disjoint, we will see that no two groups N_i and N_j with overlapping ballots can be part of one community Q . Sufficient conditions for satisfying PR when ρ consists of $\mathcal{R}_{AV}$'s only are given in the following result:

Proposition 4.6. *Let $E = (A, \mathcal{Q}, k)$. $\mathcal{R}_{RSP}^\rho$ with ρ consisting of $\mathcal{R}_{AV}$'s satisfies perfect representation if the following hold:*

- *Some voters from N_i are the majority in at least one community for all $i \in [k]$.*
- *In each community Q , for each group N_i , $\bigcap_{i \in N_i} A_{\upharpoonright Q}(i) = \{c_i^{(1)}, c_i^{(2)}, \dots, c_i^{(t)}\}$, where $c_i^{(j)}$ are the unique candidates for N_i .*
- *For all $i \neq j$, if two sub-groups $G \subseteq N_i$ and $H \subseteq N_j$ are part of the same community, $G, H \subseteq Q$, their ballots are disjoint.*

Note that it might be the case that members of one group N_i all approve of the same candidate a but one half approves of another candidate b and the other half of another candidate c . If members from the first half, that all approve of $\{a, b\}$, end up together in one community Q as a majority, $\mathcal{R}_{AV}$ elects a and b . This means that in one winner set b is elected instead of a . This is the reason why the second condition is formulated. The third condition is formulated to ensure something similar. It might be the case that members of distinct groups N_i and N_j do not share unique representatives, but they do share another candidate. In this case, we would elect the shared candidate, as her $\mathcal{R}_{AV}$ -score is higher.

Proof of theorem 4.6. Suppose that there are k groups of size $\frac{n}{k}$ that each approve of distinct candidates, that can be seen as unique representatives. The first condition ensures that for each N_i , there is a subgroup $G \subseteq N_i$ that has the highest score in at least one community. The second and third conditions guarantee that they only share unique representatives. This means that their unique candidates have the highest $\mathcal{R}_{AV}$ -score in at least one community. $\square$

Now, we consider inheritance of PR. Note that the conditions for PR are quite specific and do not reveal much about the workings of a voting rule satisfying this property in other context. For all we know, it could be possible that the rule functions as $\mathcal{R}_{AV}$ on profiles not satisfying the conditions for PR and outputs the PR result otherwise. Therefore, we consider a rule that naturally satisfies PR: Monroe's rule.

Proposition 4.7. *Let $E = (A, \mathcal{Q}, k)$ be an ABCC election. $\mathcal{R}_{RSP}^\rho$, with ρ consisting of Monroe's rule only, satisfies perfect representation if the following hold:*

- *The number of representatives from each group N_i is almost the same in each community Q . In other words, let N_i^Q denote the members of group i in community Q , for any $Q \in \mathcal{Q}$ and for any $N_i^Q, N_j^Q \subseteq Q$ it is the case that $\frac{|N_i^Q|}{|Q|} \in \{\lceil \frac{|N_j^Q|}{|Q|} \rceil, \lfloor \frac{|N_j^Q|}{|Q|} \rfloor\}$.*
- *In each community Q , for each group N_i , $\bigcap_{i \in N_i} A_{\upharpoonright Q}(i) = \{c_i^{(1)}, c_i^{(2)}, \dots, c_i^{(t)}\}$, where $c_i^{(j)}$ are the unique candidates for N_i .*

Proof. The first condition entails that the members from each group N_i are balanced in one community. As Monroe's tries to balance the number of represented voters, we get the highest Monroe score in each community by electing candidates approved by voters from distinct groups. The second condition guarantees that the unique representatives get elected. $\square$

Inheriting or satisfying perfect representation is more tedious than EJR, as it is a more demanding property. It is unlikely that the conditions imposed to inherit or satisfy PR occur naturally, as they are very specific. Do note that we did not prove it for disjoint communities. Were we to assume this, inheritance would have been easier to satisfy, as each group N_i for $i \leq k$ is contained in one community. If the weights for each community is equal to the number of different groups in that community, PR would be inherited.

4.2 Community Phragmén-like rules

In this section a new class of rules is introduced, namely the *community Phragmén-like rules*. In spirit, they will be similar to the algorithm for seq-Phragmén in Definition 3.5. Note that this algorithm does not depend on a sequence of rules, nor is there an algorithm to decide what the weights are. This makes it essentially different from the class of two-step procedures. *Inheritance*, therefore, also becomes irrelevant. To axiomatically investigate the new class of rules, some basic proportionality axioms with the aim of enforcing descriptive representation will be defined. The first axioms capture the idea that, if there is a representative that is approved by all members of a community, but does not appear in any of the other communities' profiles, this candidate should end up in the final winner set. We call such a candidate a *unique* representative.

Axiom 14 (Unique representative property (URP)). Let $E = (A, \mathcal{Q}, k)$ be an election. A winner set W satisfies the *unique representative property* if whenever there exists a $c \in C$ such that $c \in \bigcap_{i \in Q} A(i)$ for a $Q \in \mathcal{Q}$ and $c \notin \bigcup_{j \in N \setminus Q} A(i)$, then $c \in W$. A procedure $\mathcal{R}^\rho$ satisfies this procedure if for all elections $(A, \mathcal{Q}, k)$ it outputs only W 's satisfying the URP.

The following axiom is a weakening of this axiom:

Axiom 15 (Majoritarian unique representative property (MURP)). Let $E = (A, \mathcal{Q}, k)$ be an election. A winner set W satisfies the *unique representative property* if whenever there exists a $c \in C$ such that $c \in \bigcap_{i \in G} A(i)$ for a $G \subseteq Q \in \mathcal{Q}$ such that $|G| \geq \frac{|Q|}{2}$ and $c \notin \bigcup_{j \in N \setminus Q} A(i)$, then $c \in W$. A procedure $\mathcal{R}^\rho$ satisfies this procedure if for all elections $(A, \mathcal{Q}, k)$ it outputs only W 's satisfying the URP.

There is a strong similarity between this notion and perfect representation. However, these notions only coincide for very specific choices for $\mathcal{Q}$, namely those where the groups that share candidates are all part of one community. Note that $\mathcal{R}_{RSP}^\rho$ does not satisfy these axioms. Example 2.11 is a counterexample. There is a certain intuitive appeal to these axioms from the perspective of descriptive representation, as they guarantee that a community gets an unique representative. Uniqueness, in this case, can be interpreted as sharing the same characteristics as that group and understanding their particular interests. We will see, however, that it is hard to satisfy this axiom, even if we define a rule with the aims of satisfying it. This is what we do in the remainder of this section. To do this, we define the class of community Phragmén-like rules.

There will be two variants: a one-step and two-step variant. The main idea behind the general formulation is that combining budgets with voters from other communities can come with a discount, if we want to encourage intercommunal voting, or with a fine, if we want to punish intercommunal voting. I now define the one-step variant. The formulation is based on the continuous formulation of seq-Phragmén.

Definition 4.3 (One-step community Phragmén procedure (comm-Phragmén)). At $t = 0$, each voter $i \in N$ start with a budget $b_i = 0$. As t increases the budget also increases. The cost of each candidate is 1. When there is a group of voters all from one community that have a combined budget of 1 and a shared candidate, this candidate is selected. Tie-breaking can be done in several ways if there are multiple shared candidates. If, however, there are candidates from several communities, there is a fine (or discount) that increases (or decreases) the cost of a candidate. Let $d(x)$ be a general fine or (discount) function. The cost of the candidate then is $1 + d(x)$. Whenever a candidate is selected the budget for all participating voters is reset to 0.

We omit the analysis of SIOD for this procedure as it is likely that the reasoning is similar to the proof for regular $\mathcal{R}_{seqPhr}$, which can be found in the paper by Brill et al. (2023). The reason for this estimate is similar to the reasoning provided for Definition 3.5. The runtime of seq-Phragmén for the round-based formulation is determined by analysing complexity of the arithmetic operations. Note that the arithmetic operations for comm-Phragmén are essentially the same, except for the fact that the price for some groups of voters comes with a fine or discount. Determining whether this is the case is a computationally easy operation, as one only has to check if they belong to the same community.

Now, I showcase the working of comm-Phragmén with discounts and fines by applying it to Example 2.10.

Example 4.8. Let $n = 3$, $m = 4$, $k = 2$ and $\mathcal{Q} = [1, 2]$.

$$\begin{array}{l} 1 : \{a, b\} \\ \hline 2 : \{a, b\} \\ 3 : \{c, d\} \end{array}$$

We first consider a case in which we fine the voters. Define: $d : |Q| \rightarrow \mathbb{R}$ as $d(0) = 0$ and for $x > 0$, define $d(x) = \frac{1}{2}$. At $t = \frac{1}{2}$, 1 and 2 have a budget of 1, but they are fined $\frac{1}{2}$. So, they wait until $\frac{5}{8}$ and they pick a or b . At $t = 1$, 3 picks c or d .

We now consider the case with the same function, but as a discount, i.e., as $-d(x)$. Now, we see that 1 and 2 pick a or b at $t = \frac{1}{2}$, as the cost is decreased from 1 to $\frac{1}{2}$. At $t = 1$, the voters 1 and 2 and voter 3 on her own, can select a candidate. If we break ties in favour of intercommunal groups, we get $W = \{a, b\}$ otherwise $W = \{a, c\}$ or $\{a, d\}$. $\diamond$

A positive first result is that this class of rules satisfies the scaling property for any choice of fines or discounts.

Theorem 4.8. *For any choice of fines or discounts, comm-Phragmén satisfies the scaling property.*

Proof. Note that running comm-Phragmén on one community leads to the same outputs as comm-Phragmén on the whole profile.

For discounts the proof is immediate. As every voter has a voter with the same preferences in another community, any possible purchase of a candidate comes with a discount. This just means that the prize for all candidates decreases. The cost of one candidate now goes from 1 to $1 - \epsilon$ for all candidates. This means that all purchases happen earlier. Hence, this is an election in which we run seq-Phragmén on the whole profile, in which candidates are cheaper. comm-Phragmén with discounts on one profile only leads to the same result.

For fines, the same happens but it is not immediately clear. There are two options: either a fine is so high that it is cheaper to buy a candidate with fewer voters from your own community or this is not the case. In the second case, the price for all candidates increases and all purchases happen at a later time, but in the same manner. In the first case, nobody will cooperate with people from other communities as it is always cheaper to cooperate within your own community. This means that we run s community elections simultaneously, which exactly leads to the same output as the whole profile. $\square$

Note that this is a positive result. Especially as we started this chapter with the aim of finding a class of rules that satisfy the scaling property.

We now continue with the representation axioms. Note that for high fines for intercommunal cooperation, this rule is defined in such a way that it motivates voters to elect unique representatives if they exist. For this rule, we cannot derive a condition for which the URP is always satisfied. Before I present this negative result, a basic property any procedure should satisfy is presented:

Axiom 16 (Representation property (RP)). For all $Q \in \mathcal{Q}$, there exists a $c \in \bigcup_{i \in Q} A(i)$ such that $c \in W$.

Note that this is implied by the (M)URP. Unfortunately, this basic property is not satisfied by comm-Phragmén:

Proposition 4.9. *For any discount or fine, comm-Phragmén does not satisfy the RP, URP or MURP.*

Proof. Let $n = 100$, $m = 4$, $k = 4$ and $\mathcal{Q} = [5, 10, 85]$. Consider:

$$\begin{array}{r} 5 : \{a, b, c\} \\ \hline 10 : \{d, e, f\} \\ \hline 85 : \{g, h, i, j\} \end{array}$$

Note that the discount or fine does not matter, since the groups do not share any candidates. At $t = \frac{1}{85}$ g is selected, $t = \frac{2}{85}$, $\frac{3}{85}$ and $\frac{4}{85}$, respectively, h , i and j are selected. Since $\frac{4}{85} < \frac{1}{10}$, this is a correct sequence of picking candidates. Note that two communities are underrepresented. $\square$

Interesting to note about this counterexample is that each community is disjoint. For RSP, disjointness was a condition that lead to desired behaviour in terms of inheritance and satisfaction of EJR. This is not as bad as it seems, as we did not define the RP in that section. Furthermore, this profile does not satisfy the sufficient conditions we set out, as there are two communities that are too small. I now present some results for comm-Phragmén that show the interplay of this rule and the axioms we set out.

Proposition 4.10. *comm-Phragmén always satisfies the (M)URP if and only if the fine for intercommunal cooperation is greater than or equal to $n - 1 + \epsilon$ for any $\epsilon > 0$.*

Proof. We first prove the condition from right to left. Suppose that there is a fine of $n - 1 + \epsilon$. Suppose that there is an intercommunal group that wants to pick a candidate c . In the best case, each voter contributes $\frac{1}{n}$. The cost for this intercommunally liked candidate is $1 + (n - 1) + \epsilon = n + \epsilon$. This means that at time $t' = 1 + \frac{\epsilon}{n}$, this intercommunally preferred candidate will be selected. Note, however, that if there is a unique representative, she gets picked at $t = 1$.

For the other direction, suppose that the fine is lower than or equal to $n - 1$. As a counterexample, consider Example 2.9. Recall that $n = 3$, $m = 6$, $k = 4$ and $\mathcal{Q} = [1, 1, 1]$.

$$\begin{array}{c} 1 \times \{a, b, c, d\} \\ \hline 1 \times \{a, b, c, e\} \\ \hline 1 \times \{a, b, c, f\} \end{array}$$

According to the (M)URP, $\{d, e, f\}$ should be the only output. If we apply comm-Phragmén, however, at $t = 1$, any voter can pick any candidate of their approved candidates. Note that the fined candidates cost at most $1 + (3 - 1) = 3$. So, the voters either intercommunally pick anything from $\{a, b, c\}$ or any of them picks d , e or f . URP is satisfied only for one of these sets. $\square$

We see that the (M)URP is satisfied very rarely by comm-Phragmén as the fine must be very high. A further discussion of this fact is presented at the end of this chapter. Another negative result for this rule is that fining can never be combined with the satisfaction of EJR. This shows that EJR and the (M)URP are incompatible for comm-Phragmén. Note that it is not hard to see that this result can be generalised to incompatibility of EJR and RP in general.

Proposition 4.11. *comm-Phragmén satisfies EJR if and only if the fine for intercommunal cooperation is 0.*

Proof. The direction from right to left follows directly from seq-Phragmén satisfying EJR. For the other direction we reason contrapositively. Suppose that we have a fine of $\epsilon > 0$. Consider the following profile, for $n = 3$, $m = 6$, $k = 3$ and $\mathcal{Q} = [1, 1, 1]$.

$$\begin{array}{c} 1 \times \{a, b, c, d\} \\ \hline 1 \times \{a, b, c, e\} \\ \hline 1 \times \{a, b, c, f\} \end{array}$$

As intercommunal cooperation is fined, it is not cheaper to collaborate and every voter can select any candidate at $t = 1$. This means that $\{d, e, f\}$ is one of the outputs of comm-Phragmén. To satisfy EJR we have to elect $\{a, b, c\}$ only. $\square$

Lastly, we modify comm-Phragmén in such a way that it is forced to elect unique candidates. The definition of this new rule can be seen as a marriage between RSP and comm-Phragmén.

Definition 4.4 (Two-phase comm-Phragmén (2-comm-Phragmén)). In the first phase, sequentially run community elections $(A_i, 1)$ with s rules ρ for each community $i \in [s]$. Easiest is to pick $\rho = (\mathcal{R}_{AV}, \dots, \mathcal{R}_{AV})$. It is possible to break ties using a fixed tie-breaking order. After breaking ties in round i , remove the winner c from C and run $(A'_{i+1}, 1)$, note that the removed $c \in C$ does not appear in the ballots. So, we write A' instead of A . After the first phase, we end up with $C \setminus \bigcup_{i \in [s]} W_i$, where the W_i 's denotes the winner set of $\mathcal{R}_i$ applied to $(A_i, 1)$.

In the second phase, run comm-Phragmén on $E = (A^{(s)}, \mathcal{Q}, k - s)$ the whole electorate for the remaining $k - s$ seats. Here $A^{(s)}$ refers to the ballot A where the s winners from each of the community elections have been removed.

It is fairly straightforward to see that 2-comm-Phragmén satisfies URP, because of the first phase. No matter the choice of ρ , the sequential nature of the first phase guarantees that each community gets a representative. Applying it to the counterexample from the proof above, in the first phase we obtain

$\{a, d, g\}$ with breaking ties and in the second phase we obtain a candidate from $\{h, i, j\}$. Hence, a possible winner set is $\{a, d, g, h\}$. Lastly, a simple way of enforcing URP is by forcefully adding unique representatives. This can be done as follows:

- Input: an election $E = (A, \mathcal{Q}, k)$.
- Compute $\bigcap_{i \in N} A(i)$.
- For all communities $Q \in \mathcal{Q}$, make $U_Q = \bigcup_{i \in Q} A(i) \setminus \bigcap_{i \in N} A(i)$.
- For $Q \in \mathcal{Q}$ pick $c \in U_Q$ with the highest community $\mathcal{R}_{AV}$ -score and add it to W .

A further analysis of this rule is presented as part of the discussion. An important takeaway from this section is that representing communities and representing preferences are two things that are irreconcilable. If $k = s$ but the size of the communities varies significantly, there is a tension that arises from desiring representation for each community because smaller communities are given as many candidates as bigger communities. In the case of $k > s$, this problem seems less stringent as it is possible to divide the $k - s$ seats in a proportional manner.

4.3 Discussion

The following table contains the results of the rules applied to the examples from Section 2.4:

Example	comm-Phragmén with $d(x) = -\frac{1}{2}$	comm-Phragmén with $d(x) = \frac{1}{2}$	2-comm-Phragmén with $d(x) = -\frac{1}{2}$	2-comm-Phragmén with $d(x) = \frac{1}{2}$
Example 2.9	Yes	Yes	No, $\{c, d, e\}$, URP	No, $\{c, d, e\}$, URP
Example 2.10	Yes	Yes	Yes	Yes
Example 2.11	Yes	Yes	Yes	Yes
Example 2.12	Yes	No, $\{c, d\}$	Yes	Yes
Example 2.13	Yes	No, $\{c, d, e\}$, URP	Yes	No, $\{c, d, e\}$, URP
Example 2.14	Yes	No, $\{c, d, e\}$, URP	No, $\{c, d, e\}$, URP	No, $\{c, d, e\}$, URP
Example 2.15 (Fatal flaw)	Yes	Yes	No, $\{a, b\}$	No, $\{a, b\}$
Example 2.16	Yes	Yes	Yes	Yes
Example 2.17	Yes	Yes	Yes	Yes
Example 2.18	Yes	Yes	Yes	Yes
Example 2.20	Yes	Yes	No, $\{a, b, c, d\}$, URP	No, $\{a, b, c, d\}$, URP

Table 4.1: This table contains whether the community Phragmén rules satisfy EJR for the examples from Section 2.4. If EJR is not satisfied, a winner set is stated that does not satisfy EJR. If URP is satisfied for this set, this is mentioned.

It would be interesting to see for which discounts d comm-Phragmén with discounts satisfies EJR. It is likely that it is not the case for all discounts d , but there is no obvious argument that excludes the satisfaction of EJR for smaller d . comm-Phragmén with fines and 2-comm-Phragmén with fines and discounts do not always satisfy EJR. This is something one can expect, as the first phase of these

procedures is non-proportional and, for these examples, the second phase does not exist as $k = s$ or it is for a small number of remaining candidates. Adding a fine decreases the number of candidates that are liked by voters from different communities.

This section contains some basic ideas on how to use the community structure without fixing seats for each community. The two rules that were designed show that it is not possible to satisfy the URP and EJR. Which points towards the incompatibility of rigid forms of descriptive and substantive representation. Insisting on the satisfaction of the (M)URP can lead to a form of sectarianism, which, if it occurs, is seen as a negative societal phenomenon. See Bishara (2021) and Weiss (2015) for in-depth philosophical discussion of this phenomenon. The (M)URP can be seen as a rule that enforces a form of sectarianism. As it prefers candidates that have high local support over candidates that have higher global support. Consider:

$$\frac{1 \times \{a, b, c, d\}}{1 \times \{a, b, c, e\}} \\ \frac{1 \times \{a, b, c, e\}}{1 \times \{a, b, c, f\}}$$

In this example $\{d, e, f\}$ is preferred over $\{a, b, c\}$, which means that a community-only representatives is preferred to global representatives, even if they are approved by everyone. As mentioned before, if $k > s$, every community should get a representatives, but if $k = s$ a certain part of the electorate should be voted on by everyone as this situation is too susceptible for forms of sectarianism.

This observation is a plea for a kind of middle-way in between justified representation and community representation. Theorems 4.10 and 4.11 offer quantitative perspectives on the RP. We see that to always satisfy the (M)URP we need to force voters to *never* cooperate with any other candidate; even a unanimously approved candidate must be neglected. EJR, on the other hand, is always satisfied if descriptive representation is neglected. In hindsight, the latter half of this chapter was devoted to investigating the extremes of the spectrum between descriptive representation (URP) and substantive representation (EJR). comm-Phragmén for mild fines and discounts can shine light on the parts in between. Investigating comm-Phragmén for different choices of fines and discounts seems to be an interesting endeavour if we wish to formally understand descriptive representation and its interplay with substantive forms of representation.

Conclusions

The main question we set out to answer was: *Is it possible to represent voters in the ABC model? If so, in how far is it possible to combine this form of representation with substantive forms of representation, in particular, EJR and PR?*

Put differently, is it possible to balance descriptive and substantive representation in ABC voting? The short, and perhaps frustrating, answer is that it depends. If we view compatibility between the two notions of representation in terms of inheritance of proportionality axioms by RSP — which was defined to enforce descriptive representation — it is clearly impossible to balance the two. On the other hand, we saw that if all members of each community collectively approve of at least one candidate, it is possible to have a representative for each community, while simultaneously satisfying justified representation. We also saw that if each community had disjoint ballots and approved of disjoint parts of the set of candidates, the inheritance and satisfaction of justified representation is guaranteed under relatively undemanding conditions.

In general, it is not possible to represent voters while simultaneously trying to fairly divide the total utility given the preferences of those voters. This sheds some light on the distinction between descriptive and substantive representation: in cases where belonging to a group seems to be independent from that group's preferences, descriptive and substantive representation are incompatible. Intuitively, this also makes sense. If a community with a shared demographic trait has distinct and opposing preferences, it is unlikely that any one representative from that group will be able to work in all of their interests. This intuition is substantiated by the non-inheritance of all proportionality axioms and the incompatibility of the representation property (RP) with both extended justified representation (EJR) and perfect representation (PR).

This thesis can be seen as a different approach to one of the problems Kagita et al. (2021) tackle in their work. Recall that they generalise the framework of ABC voting by making candidates tuples of attributes which voters can approve of. With the application of descriptive voter representation in mind, the attributes can be taken to be demographic traits. However, on a high level, it could be that voters approve of demographic traits that they themselves do not share. Hence, it could be the case that if we represent voters based on their attribute preferences, they do not get represented descriptively. The ABCC model takes a similar approach, but attributes are assigned to voters and voters belonging to one group get a greater say in the determination of the final winner set. The conclusion is that the ABCC voting procedures, just like the AAV voting rules, are unable to satisfy basic (substantive) proportionality axioms.

Results per chapter

Chapter 2 This chapter served as a motivation for descriptive representation. We argued that, theoretically speaking, substantive representation is a more meaningful form of representation, but in many cases it is not possible to have the substantive kind without the descriptive kind. Afterwards, we discussed potential ways of modelling demographic traits. Instead of opting for a model in which demographic traits are added explicitly, we chose for an implicit approach in which the only information that was encoded as part of the model was that of voters belonging to communities. This was done by partitioning the set of voters N into s sets. Each set corresponded to a community and elements of this set are seen as community members.

Next to this, we defined two classes of voting rules, based on different considerations. One, called the *reserved seats procedure (RSP)* and the other was called the *naively weighted procedure (NWP)*. The idea behind the first rule is that each community gets a fixed number of candidates in the final winner set they vote for. These are then filled by the winners of that community. The main problems of this voting rule were determining the weights for each of the communities and combining intersecting community winner sets into a final winner set. NWP was not subjected to these problems. This procedure first runs elections inside each community for a winner set of size k and then aggregates the community winner sets into one final winner set by using another ABC rule. These functions were implemented into Python and their satisfaction of EJR was investigated for small n , m and k and for all possible ways of dividing n into communities.

The goal of the systematic generation was to find all profiles for which EJR was not satisfied. The main findings were that NWP performed worse than RSP in terms of satisfaction of EJR. There were more combinations of variables for which it did not satisfy EJR and the number of elections for which it did not satisfy EJR was higher. The reason for why RSP did not satisfy EJR was that the ‘rescaling’ of the final winner set lead to the (dis)appearance of cohesive groups. A group that is ‘big enough’ with respect to the full profile, might get split up into distinct communities and lose their cohesiveness. The biggest problems of NWP were similar. It functioned well when communities were cohesive, but when group were not cohesive it led to results that were less proportional than the least proportional voting rules and it also led to a huge jump in runtime. In general, it became clear that it is impossible to simultaneously guarantee extreme forms of descriptive and substantive representation.

Chapter 3 Whereas the second chapter corresponded to work done mainly in the pre-axiomatic phase of the history of social choice theory, Chapter 3 was devoted to axiomatic and computational questions. We investigated the axiomatic and computational properties of the defined voting rules. The main findings were that RSP performed better in terms of *inheritance* of ABC axioms than NWP. NWP inherited three axioms: strategyproofness and support monotonicity with and without additional voters. Out of the most well-known natural voting rules, only AV satisfies these axioms. An additional problem is that $\mathcal{R}_{NWP}^\rho$ with ρ consisting of all- $\mathcal{R}_{AV}$ ’s led to particularly bad results in terms of proportionality, which we saw in Section 2.4.

RSP inherited all axioms that NWP inherited and two more: committee monotonicity and internal consistency from consistent axioms. There are multiple voting rules satisfying some of these axioms, making RSP better behaved and a relatively appealing voting procedure. Big downsides of both rules is their non-inheritance of weak efficiency and the intractability of determining whether a given subset is part of their final output. Weak efficiency is a rudimentary axiom any sensible voting procedure should

satisfy, as it captures the idea that approved candidates should always be elected over non-approved candidates. Fortunately, the original formulations of RSP and NWP can be modified in such a way that they do satisfy weak efficiency. The intractability, however, is a deeper problem that makes both of these procedures unsuited for large-scale applications. This is a negative result, as a primary motivator for the formulation of these rules are parliamentary elections.

For two-step procedures, inheritance of axioms is an important indication of their performance. The reason being that inheritance of an axiom is a weaker property than satisfaction. If a voting procedure $\mathcal{R}^\rho$ satisfies an axiom for any election $E = (A, \mathcal{Q}, k)$ and choice of voting rules ρ , it also inherits that property P from voting rules satisfying P . Non-inheritance in the context of ABCC voting is just as bad as non-satisfaction of a property in the context of ABC voting. Therefore, the non-inheritance of all proportionality axioms can be seen as a negative result. In particular, NWP and RSP performed poorly if communities shared a lot of candidates amongst each other. Using rules that are supposed to represent communities with similar preferences, in the context of communities that do not have similar preferences, leads to unsatisfactory results in terms of proportionality. The observations around these axioms paved the way for the discussion of Chapter 4.

Chapter 4 The title of this chapter is *escape routes*, which were set out to escape from the problems that occurred when analysing proportionality for RSP and NWP. Three possible escape routes were identified: domain restrictions, new voting rules and new axioms. A particular problem that was identified was the so-called *fatal flaw*:

1 : $\{a, b\}$
2 : $\{c, d\}$
3 : $\{a, b\}$
4 : $\{c, d\}$

This was dubbed the fatal flaw because it was argued that intuitively speaking, there is no proportional (either descriptive or substantive) reason to choose $\{a, b\}$ and $\{c, d\}$. This showed that voting rules and axioms should pick something from $\{a, b\} \times \{c, d\}$ for this election. Unfortunately, both $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ returned these sets. An axiom that was defined was the *scaling property*. This captures the idea that, if it is possible to scale the larger election in such a way that the s smaller elections are scaled variants of this larger election, the outcome of applying the larger procedure to one community should be identical to applying the procedure to the whole election. It was shown that RSP did not satisfy it for ρ consisting of standard ABC rules only. This, again, was a negative result for RSP.

Another class of ABCC axioms that were defined were representation properties. These captured the idea that each community should get at least one representative, given certain conditions. We defined three variants: the representation property (RP), the unique RP (URP) and the majoritarian URP (MURP). Note that RSP satisfied RP, tfor any diverse set of weights $\mathbf{q}$. It also became immediately apparent that these kinds of representation axioms are incompatible with EJR and PR. This was another result in a long list pointing towards the incompatibility of representing voters and representing voter preferences.

With the incompatibility of descriptive and substantive representation in mind, a new class of voting rules was defined: *community Phragmén-like voting rules*. The main idea between this class

of voting rules and RSP, was that representation was not guaranteed for all communities. In fact, it allows for cooperation between different communities by fixing parts of W . These rules functioned like the regular sequential version of Phragmén’s voting method, but electing a candidate with someone from a different community came with a fine (or discount). The idea being that, this way, cooperation with your own community is not *enforced*, but *supported*. The downside of a ‘soft’ rule, in terms of guarantee of a descriptive representative, is that it does not satisfy any of the representation properties for a reasonable fine. Lastly, we presented a rule that consisted of two phases. It, first, guaranteed representation for every group and then it ran comm-Phragmén for the remainder of W . This way, there is a ‘balance’ of having descriptive and substantive representatives.

The domain restrictions came with some positive results: when the communities are disjoint, i.e., each community votes on different candidates, JR and EJR are under mild conditions. Determining a winner in this restricted domain can be done in polynomial time. Using ρ with all- $\mathcal{R}_{AV}$ ’s led to the satisfaction of JR if each community shares one candidate. Note that this is one way in which the ABCC model can guarantee a form of proportionality *and* strategyproofness. Hence, using RSP when communities are disjoint is sensible. For comm-Phragmén, this was the domain for which it performed worst. In the context where communities are not disjoint and there is much overlap between communities, formulating conditions that lead to EJR or JR came with more difficulties. The same can be said for PR. This substantiated the idea that working with fixed seats makes sense only if the electorates are cohesive and distinct. The investigation of comm-Phragmén showed that EJR and representation properties are extremes on a scale. This pointed to the need of understanding the grey areas between these extremes.

Future work

Lastly, I discuss possible directions for further formal research on descriptive representation. This is divided into three general categories: connecting to other formalisms, impossibilities and axioms, and practical questions.

There is a natural connection to apportionment. This can be seen as a restriction of the domain of approval-based voting. Instead of approving candidates as individuals, voters are restricted to voting on one party-list. Note that the case in which all communities are disjoint can be seen as a ‘disjoint sum’ of elections. Hence, if a rule is proportional for local elections it is proportional with respect to the whole election. Interesting to investigate is the case in which communities are not disjoint. An important question is: does restricting to party-lists lead to milder conditions for the satisfaction of proportionality axioms? Furthermore, if we assume $\mathcal{Q}$ to be disjoint we can simulate historical elections that took place in colonial India before 1947.

Another related formalism is the *approval-based apportionment* framework developed by Brill et al. (2022). This is a framework that can be placed between apportionment and approval voting. In this context, voters can approve of multiple parties. Next to the conceptual similarity of this framework to the intended application of the ABCC model, restricting to approval-based apportionment comes with a computational advantage: PAV always returns a core-stable committee, a strong proportionality axiom, and it runs in polynomial time. If it can be shown that this restricted domain does lead to a certain ABCC rule inheriting or satisfying EJR, while simultaneously guaranteeing a form of descriptive representation, we would have achieved the goal of balancing descriptive with substantive representation and tractability.

Although this thesis did include some basic and simple impossibility theorems, such as the incompatibility of RP and EJR, it would be interesting to explore this realm further to better understand the ABCC procedures. An important question regarding this is the inheritance of proportionality axioms by two-step procedures. Is it the case that it is impossible to inherit EJR given certain restrictions on the two steps, i.e., on the tuple ρ , the determination of the weights $\mathbf{q}$ and restrictions on the determination of the final set W ? Or does there exist a two-step procedure that inherits certain proportionality axioms for all choices of $\mathcal{Q}$? Next to this, the community Phragmén-like rules can be examined further. Which fines lead to a sensible balance between descriptive and substantive representation? Is it possible to define axioms that capture a form of descriptive representation that are compatible with substantive representation? And, what is the runtime of these community Phragmén-like rules?

Some unanswered questions that arose during the axiomatic analysis in Chapter 3 were relegated to the future. I now provide an overview of these questions. First, is it the case that if we modify RSP and NWP to satisfy weak efficiency ($\uparrow$) the same axioms are inherited as for the original formulation? Furthermore, support monotonicity *without* extra voters is inherited. Does the same hold if we do add extra voters? Or an extra community ($\uparrow$)? Lastly, recall that consistency cannot be directly lifted to the ABCC context. The ‘external’ variant that was defined is rather artificial and not inherited ($\uparrow$). Does there exist a natural ABCC variant of consistency that is weaker?

The motivation for the definition of the ABCC model and the definition of the voting rules were intended for a specific practical purpose. This application, unfortunately, has not been a part of this thesis. It would be interesting to see how the defined rules perform on datasets that can naturally capture cohesive to non-cohesive community structures. In particular, as many counterexamples we encountered did not correspond to natural situations; for instance, $|\mathcal{Q}| = 3$ and $n = 3$. Restricting the choice of $\mathcal{Q}$ based on real life observation would be interesting. A natural choice is $\mathcal{Q}$ with two roughly equal communities and one where there is a majority and minority community. For parliamentary elections this would correspond to the demographic make-up of many countries.

On the whole, I hope to have contributed to a better understanding of descriptive representation and its relation to substantive representation. The method employed for this was computational and axiomatic, borrowing solely from computer science and social choice theory. As far as my knowledge reaches, such a formal analysis of descriptive representation has not been done before. Next to the technical results and goals, one of the aims of this thesis was to connect two disjoint perspectives on representation; one from the perspective of political philosophy and the other developed in the context of computational social choice.

Bibliography

- Arora, Sanjeev and Boaz Barak (2009). *Computational Complexity: A Modern Approach*. 1st edition. Cambridge University Press (cited on page 51).
- Arrow, Kenneth J. (1951). *Social Choice and Individual Values*. 2nd edition. John Wiley and Sons (cited on page 4).
- Arrow, Kenneth J., Amartya K. Sen, and Kotaro Suzumura, editors (2002). *Handbook of Social Choice and Welfare*. Volume 1. Elsevier (cited on page 3).
- Aziz, Haris, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh (2017). “Justified representation in approval-based committee voting”. In: *Social Choice and Welfare* 48, pages 461–485 (cited on pages 2, 5, 7, 13, 15, 16).
- Aziz, Haris, Edith Elkind, Shenwei Huang, Martin Lackner, Luis Sánchez-Fernández, and Piotr Skowron (2018). “On the complexity of extended and proportional justified representation”. In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-2018)*. AAAI Press, pages 902–909. ISBN: 978-1-57735-800-8 (cited on page 14).
- Aziz, Haris, Serge Gaspers, Joachim Gudmundsson, Mackenzie Simon, Nicholas Mattei, and Toby Wals (2015). “Computational Aspects of Multi-Winner Approval Voting”. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent System (AAMAS-2015)*, pages 107–115 (cited on page 51).
- Balinski, Michel L. and Peyton H. Young (2001). *Fair Representation Meeting the Ideal of One Man, One Vote*. Volume 2. Brookings Institution Press (cited on page 17).
- Barberà, Salvador (1977). “Manipulation of social decision functions”. In: *Journal of Economic Theory* 15.2, pages 266–278 (cited on page 49).
- Barberà, Salvador and Prasanta K. Pattanaik (1984). “Extending an order on a set to the power set: some remarks on Kannai and Peleg’s approach.” In: *Journal of Economic Theory* 32.1, pages 185–191 (cited on page 49).
- Bird, Karen (2010). “Patterns of substantive representation among visible minority MPs: Evidence from Canada’s House of Commons”. In: *The Political Representation of Immigrants and Minorities*. Edited by Karen Bird, Thomas Saalfeld, and Andreas M. Wüst. Routledge. Chapter 9 (cited on page 20).
- Bird, Karen, Thomas Saalfeld, and Andreas M. Wüst, editors (2010). *The Political Representation of Immigrants and Minorities*. Routledge (cited on pages 2, 20).
- Bishara, Azmi (2021). *Sectarianism without Sects*. Oxford University Press (cited on page 71).
- Bogaards, Matthijs (2019). “Consociationalism and Centripetalism: Friends or Foes?” In: *Swiss Political Science Review* 25.4, pages 519–537 (cited on page 19).
- Botan, Sirin (2021). “Strategyproof Social Choice for Restricted Domains”. PhD thesis. ILLC, University of Amsterdam (cited on pages 49, 50).

- Brams, Steven J. and Peter C. Fishburn (2010). “Going from Theory to Practice: The Mixed Success of Approval Voting”. In: *Handbook on Approval Voting*. Edited by Jean-François Laslier and M. Remzi Sanver. Studies in Choice and Welfare. Springer International Publishing. Chapter 3 (cited on page 5).
- Brandt, Felix, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors (2016a). *Handbook of Computational Social Choice*. Cambridge University Press (cited on pages 2, 3).
- (2016b). “Introduction to the Theory of Voting”. In: *Handbook of Computational Social Choice*. Edited by Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Cambridge University Press. Chapter 1 (cited on page 4).
- Brill, Markus, Rupert Freeman, Svante Janson, and Martin Lackner (2023). “Phragmén’s voting methods and justified representation”. In: *Mathematical Programming* 203.1-2, pages 47–76 (cited on pages 12, 47, 52, 67).
- Brill, Markus, Paul Gözl, Dominik Peters, Ulrike Schmidt-Kraepelin, and Kai Wilker (2022). “Approval-based apportionment”. In: *Mathematical Programming* 203, pages 77–105 (cited on page 75).
- Campbell, Donald E. and Jerry S. Kelly (2002). “Impossibility Theorems in the Arrovian Framework”. In: *Handbook of Social Choice and Welfare*. Edited by Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura. Volume 1. Elsevier. Chapter 1 (cited on page 4).
- Casellas, Jason P. and David L. Leal (2010). “Minority representation in the US Congress”. In: *The Political Representation of Immigrants and Minorities*. Edited by Karen Bird, Thomas Saalfeld, and Andreas M. Wüst. Routledge. Chapter 8 (cited on page 20).
- Chamberlin, John R. and Paul N. Courant (1983). “Representative Deliberations and Representative Decisions: Proportional Representation and the Borda Rule”. In: *American Political Science Review* 77.3, pages 718–733 (cited on page 10).
- Conitzer, Vincent and Toby Walsh (2016). “Barriers to Manipulation in Voting”. In: *Handbook of Computational Social Choice*. Edited by Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Cambridge University Press (cited on page 49).
- Elkind, Edith, Piotr Faliszewski, Ayumi Igarashi, Pasin Manurangsi, Ulrike Schmidt-Kraepelin, and Warut Suksompong (2024). “The Price of Justified Representation”. In: *ACM Transactions on Economics and Computation* 12.3, pages 1–27 (cited on page 16).
- Elkind, Edith, Piotr Faliszewski, Piotr Skorwon, and Arkadii Slinko (2017). “Properties of multiwinner voting rules”. In: *Social Choice and Welfare* 48, pages 599–632 (cited on pages 34, 38).
- Faliszewski, Piotr and Piotr Skowron (2017). “Chamberlin–Courant Rule with Approval Ballots: Approximating the MaxCover Problem with Bounded Frequencies in FPT Time”. In: *Journal of Artificial Intelligence Research* 60, pages 687–716 (cited on page 54).
- Fishburn, Peter C. (1972). “Even-chance lotteries in social choice theory”. In: *Theory and Decision* 3.1, pages 18–40 (cited on page 49).
- (1973). *The Theory of Social Choice*. Princeton University Press (cited on page 3).
- Gärdenfors, Peter (1976). “Manipulation of Social Choice Functions”. In: *Journal of Economic Theory* 13.2, pages 217–228 (cited on page 49).
- Geist, Christian and Ulle Endriss (2011). “Automated Search for Impossibility Theorems in Social Choice Theory: Ranking Sets of Objects”. In: *Journal of Artificial Intelligence Research* 40, pages 143–174 (cited on page 4).

- Gibbard, Allan (1973). “Manipulation of Voting Schemes: A General Result”. In: *Econometrica* 41.4, pages 587–601 (cited on pages 4, 49).
- Goel, Ashish, Anilesh K. Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto (2019). “Knapsack Voting for Participatory Budgeting”. In: *ACM Transactions on Economics and Computation* 7.2, pages 1–27 (cited on page 28).
- Haye, Freek (Jan. 28, 2025). “Opgeknapte plaskrul voor minimaal 1500 euro aangeboden op Marktplaats: ‘Hij is helemaal gerestaureerd’”. In: *Het Parool*. <https://www.parool.nl/amsterdam/opgeknapte-plaskrul-voor-minimaal-1500-euro-aangeboden-op-marktplaats-hij-is-helemaal-gerestaureerd~b4d1b0435> (cited on page 6).
- Kagita, Venkateswara Rao, Arun K. Pujari, Vineet Padmanabhan, Haris Aziz, and Vikas Kumar (2021). “Committee Selection using Attribute Approvals”. In: *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS-2021)*, pages 683–691 (cited on pages 7, 19, 21–23, 72).
- Kannai, Yakar and Bezalel Peleg (1984). “A note on the extension of an order on a set to the power set”. In: *Journal of Economic Theory* 31.1, pages 172–175 (cited on page 49).
- Karp, Richard M. (1972). “Reducibility Among Combinatorial Problems”. In: *Complexity of Computer Computations*. Edited by R.E. Miller, Thatcher J.W., and J.D. Bohlinger. Plenum, pages 185–191 (cited on page 53).
- Kelly, Jerry S. (1977). “Strategy-Proofness and Social Choice Functions without Singlevaluedness”. In: *Econometrica* 45.2, pages 439–446 (cited on page 49).
- Kiesraad (2021). *Verkiezingsuitslag Tweede Kamer 17 maart 2021*. <https://www.verkiezingsuitslagen.nl/verkiezingen/detail/TK20210317>. Accessed: 09-06-2025 (cited on page 4).
- Kluiving, Boas, Adriaan de Vries, Pepijn Vrijbergen, Arthur Boixel, and Ulle Endriss (Aug. 2020). “Analysing Irresolute Multiwinner Voting Rules with Approval Ballots via SAT Solving”. In: *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI-2020)* (cited on page 36).
- Lackner, Martin and Piotr Skowron (2020a). “Consistent approval-based multi-winner rules”. In: *Journal of Economic Theory* 192.105173 (cited on pages 36, 37, 44–46).
- (2020b). “Utilitarian welfare and representation guarantees of approval-based multiwinner rules”. In: *Artificial Intelligence* 288.103366 (cited on pages 14, 37, 38).
- (2023). *Multi-Winner Voting with Approval Preferences*. SpringerBriefs in Intelligent Systems. Springer International Publishing (cited on pages 5, 7, 9, 10, 13, 19, 34, 39, 41, 47, 50, 55).
- Lang, Jérôme and Piotr Skowron (2018). “Multi-Attribute Proportional Representation”. In: *Artificial Intelligence* 263, pages 74–106 (cited on page 7).
- Laslier, Jean-François and M. Remzi Sanver, editors (2010). *Handbook on Approval Voting*. Studies in Choice and Welfare. Springer International Publishing (cited on page 5).
- Lijphart, Arend (1969). “Consociational Democracy”. In: *World Politics* 21.2, pages 207–225 (cited on pages 8, 19).
- Mansbridge, Jane (1999). “Should Blacks Represent Blacks and Women Represent Women? A Contingent “Yes””. In: *The Journal of Politics* 61, pages 628–657 (cited on pages 7, 20).
- Monroe, Burt L. (1995). “Fully Proportional Representation”. In: *The American Political Science Review* 89.4, pages 925–940 (cited on pages 7, 12).

- Oudhof, Ko, Carel Harmsen, Suzanne Loozen, and Chan Choenn (2011). “Omvang en spreiding van Surinaamse bevolkingsgroepen in Nederland”. In: (cited on page 23).
- Ozkes, Ali I. and M. Remzi Sanver (2021). “Anonymous, neutral, and resolute social choice revisited”. In: *Social Choice and Welfare* 57, pages 97–113 (cited on page 35).
- Peters, Dominik (2018). “Proportionality and strategyproofness in multiwinner elections”. In: *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2018)*, pages 1549–1557 (cited on pages 49, 50).
- (2019). “Fair Division of the Commons”. PhD thesis. University of Oxford (cited on pages 36, 49).
- Peters, Dominik and Piotr Skowron (2020). “Proportionality and the limits of welfarism”. In: *Proceedings of the 2020 ACM Conference on Economics and Computation (ACM-EC-2020)*. Extended and updated version can be found here: <https://arxiv.org/abs/1911.11747>, pages 793–794 (cited on pages 2, 7, 12, 46, 47).
- Pettit, Philip (2012). “Varieties of public representation”. In: *Political Representation*. Edited by Ian Shapiro, Susan C. Stokes, Elisabeth Jean Wood, and Alexander S. Kirshner. Cambridge University Press (cited on pages 2, 7, 19).
- Procaccia, Ariel D., Jeffrey S. Rosenschein, and Aviv Zohar (2008). “On the Complexity of Achieving Proportional Representation”. In: *Social Choice and Welfare* 30.3, pages 353–352 (cited on page 51).
- Pukelsheim, Friedrich (2017). *Proportional Representation. Apportionment Methods and Their Applications*. 2nd edition. Springer International Publishing (cited on pages 7, 16, 17).
- Rai, Shirin M. (2005). “Reserved Seats in South Asia: A Regional Perspective”. In: *Women in Parliament: Beyond Numbers*. Edited by Julie Ballington and Azza Karam. International IDEA, pages 174–184 (cited on pages 8, 19).
- Reitzig, Raphael and Sebastian Wild (2024). “A simple and fast linear-time algorithm for divisor methods of apportionment”. In: *Mathematical Programming* 203, pages 187–205 (cited on page 17).
- Ruedin, Didier (2020). “Regional and Ethnic Minorities”. In: *The Oxford Handbook of Political Representation in Liberal Democracies*. Edited by Robert Rohrschneider and Jacques Thomassen. Oxford University Press (cited on pages 2, 5, 7, 19, 20).
- Sánchez-Fernández, Luis, Edith Elkind, Martin Lackner, Norberto Fernández, Jesús A. Fisteus, Pablo Basanta Val, and Piotr Skowron (2017). “Proportional Justified Representation”. In: *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*, pages 670–676 (cited on pages 2, 7, 15, 29, 47).
- Sánchez-Fernández, Luis and Jesús A. Fisteus (2018). “Monotonicity Axioms in Approval-based Multiwinner Voting Rules”. In: *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAAMAS-2019)*, pages 485–493. ISBN: 978-1-57735-800-8 (cited on pages 40, 41).
- Satterthwaite, Mark Allen (1975). “Strategy-proofness and Arrow’s conditions. Existence and correspondence theorems for voting procedures and social welfare functions”. In: *Journal of Economic Theory* 10.2, pages 187–217 (cited on pages 4, 49).
- Skowron, Piotr (2021). “Proportionality degree of multiwinner rules”. In: *Proceedings of the 2021 ACM Conference on Economics and Computation (ACM-EC-2021)*, pages 820–821 (cited on page 14).
- Skowron, Piotr, Piotr Faliszewski, and Jérôme Lang (2016). “Finding a Collective Set of Items: From Proportional Multirepresentation to Group Recommendation”. In: *Artificial Intelligence* 241, pages 191–216 (cited on page 51).

- Smith, John H. (1973). “Aggregation of Preferences with Variable Electorate”. In: *Econometrica* 41.6, pages 1027–1041 (cited on page 44).
- Suzumura, Kotaro (2002). “Introduction”. In: *Handbook of Social Choice and Welfare*. Edited by Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura. Volume 1. Elsevier (cited on page 3).
- Tan, Netina and Cassandra Preece (2022). “Ethnic Quotas, Political Representation and Equity in Asia Pacific”. In: *Journal of Representative Democracy* 58.3, pages 347–371 (cited on pages 8, 19).
- Thiele, Thorvald N. (1895). “Om flerfoldsvalg”. In: *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441 (cited on page 10).
- Weiss, Max (2015). “The Matter of Sectarianism”. In: *The Oxford Handbook of Contemporary Middle Eastern and North African History*. Edited by Amal Ghazal and Jens Hanssen. Oxford University Press. Chapter 8, pages 144–161 (cited on page 71).
- Young, H. Peyton (1974). “A Note on Preference Aggregation”. In: *Econometrica* 42.6, pages 1129–1131 (cited on page 44).
- (1975). “Social Choice Scoring Functions”. In: *SIAM Journal on Applied Mathematics* 28.4, pages 824–838 (cited on page 44).
- Zwicker, William S. (2016). “Introduction to the Theory of Voting”. In: *Handbook of Computational Social Choice*. Edited by Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Cambridge University Press (cited on page 3).

Appendix A

Implementation

This appendix contains the formalisation that was used for the implementation of the two voting procedures in Python. The second section contains the quantitative analysis of both applied to all profiles for small n , m and k , and for all possible ways of dividing n into communities. The focus is understanding the satisfaction of EJR.

A.1 RSP and NWP as algorithms

This section contains the $\mathcal{R}_{RSP}^\rho$ and $\mathcal{R}_{NWP}^\rho$ in pseudo-code, corresponding to how they were implemented in Python.

Algorithm 1 Naively Weighted Procedure (NWP)

Require: ABCC election $(A, \mathcal{Q}, k)$, voting rules $(R_1, \dots, R_s)$, special rule R_{s+1} .

Ensure: Set K of winner sets, each of size k .

```
1:  $\mathbf{q} \leftarrow (|N_1|, |N_2|, \dots, |N_s|)$ 
2: for  $i \in [s]$  do
3:    $R_i \leftarrow R_i(A_i, k)$ 
4: end for
5:  $X \leftarrow \times_{i \in [s]} R_i$ 
6:  $\mathbf{A} \leftarrow \emptyset$ 
7: for all  $x \in X$  do
8:    $\Pi \leftarrow (0)_{i \leq n}$ 
9:   for  $i = 1, \dots, s$  do
10:    for  $j = 1, 2, \dots, q_i$  do
11:       $\Pi[j] = x[i]$ 
12:    end for
13:  end for
14:   $\mathbf{A} \leftarrow \mathbf{A} \cup \{\Pi\}$ 
15: end for
16:  $K \leftarrow \emptyset$ 
17: for all  $\Pi \in \mathbf{A}$  do
18:    $K \leftarrow K \cup R_{s+1}(\Pi, k)$ 
19: end for
20: return  $K$ 
```

Algorithm 2 Reserved Seats Procedure (RSP)

Require: ABCC election $(A, \mathcal{Q}, k)$, voting rules $\rho = (\rho_1, \dots, \rho_s)$, apportionment method F .

Ensure: Set K or L of winner sets, each of size k .

```
1:  $q \leftarrow F(|N_1|, \dots, |N_s|), k)$ 
2: for  $i = 1, \dots, s$  do
3:    $R_i \leftarrow \mathcal{R}_i(A_i, q_i)$ 
4: end for
5:  $X \leftarrow \times_{i \in [s]} R_i$ 
6:  $K, L \leftarrow \emptyset$ 
7: for all  $x \in X$  do
8:    $U \leftarrow \bigcup x$ 
9:   if  $|U| = k$  then
10:     $K \leftarrow K \cup \{U\}$ 
11:   else
12:     $L \leftarrow L \cup \{U\}$ 
13:   end if
14: end for
15: if  $K \neq \emptyset$  then
16:   return  $K$ 
17: else
18:   for all  $U \in L$  do
19:     for all  $Y \subseteq (C \setminus U)$  such that  $|U \cup Y| = k$  do
20:        $S \leftarrow S \cup \{U \cup Y\}$ 
21:     end for
22:   end for
23:   return  $L$ 
24: end if
```

A.2 Results of the systematic generation of elections

The profiles were generated for all $m \leq 6$ until $n \leq 9$ and for all $k \in [1, m]$. The results are given in the tables below. If EJR is not satisfied, (A, B) is written in the cell. A denotes the percentage of profiles for which EJR is not satisfied and B shows the likelihood of randomly picking a committee for which EJR is not satisfied in such a profile. For $n \in [5, 8]$, m is quite small. We see that for small m it is easier to satisfy EJR.

n	m	k	EJR for RSP	Partitions	EJR for NWP	Partitions
3	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	
3	4	1	Yes		Yes	
3	4	2	Yes		No	$\{\{1\}, \{2, 3\}\}(0.056, 0.167)$
3	4	$\{3, 4\}$	Yes		Yes	
3	5	1	Yes		Yes	
3	5	2	Yes		No	$\{\{1\}, \{2, 3\}\}(0.18, 0.167)$
3	5	3	No	$\{\{1\}, \{2, 3\}\}(0.12, 0.167)$ $\{\{1\}, \{2\}, \{3\}\}(0.6, 0.2)$	No	$\{\{1\}, \{2, 3\}\}(0.18, 0.167)$
3	5	$\{4, 5\}$	Yes		Yes	
3	6	1	Yes		Yes	
3	6	2	Yes		No	$\{\{1\}, \{2, 3\}\}(0.267, 0.167)$
3	6	3	No	$\{\{1\}, \{2, 3\}\}(0.18, 0.22)$ $\{\{1\}, \{2\}, \{3\}\}(0.72, 0.2)$	No	$\{\{1\}, \{2, 3\}\}(0.46, 1)$ $\{\{1\}, \{2\}, \{3\}\}(0.01, 1)$
3	6	4	Yes	$\{\{1\}, \{2\}, \{3\}\}(0.04, 0.2)$	No	
3	6	5	Yes		Yes	
4	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	
4	4	$\{1, 3, 4\}$	Yes		Yes	
4	4	2	No	$\{\{1, 2\}, \{3, 4\}\}(0.01, 0.33)$	No	$\{\{1, 2\}, \{3, 4\}\}(0.01, 0.33)$ $\{\{1\}, \{2, 3, 4\}\}(0.014, 1)$
4	5	1	Yes		Yes	
4	5	2	No	$\{\{1, 2\}, \{3, 4\}\}(0.08, 0.4)$		$\{\{1, 2\}, \{3, 4\}\}(0.66, 0.37)$ $\{\{1\}, \{2, 3, 4\}\}(0.12, 1)$
4	5	3	No	$\{\{1, 2\}, \{3, 4\}\}(0.01, 0.17)$ $\{\{1\}, \{2, 3, 4\}\}(0.02, 0.11)$	No	$\{\{1, 2\}, \{3, 4\}\}(0.11, 0.07)$ $\{\{1\}, \{2\}, \{3, 4\}\}(0.05, 0.1)$
4	5	4	Yes			Yes
5	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	
5	4	1	Yes		Yes	
5	4	2	Yes		No	$\{\{1\}, \{2, 3, 4, 5\}\}(0.005, 0.167)$ $\{\{1, 2\}, \{3, 4, 5\}\}(0.002, 1)$
6	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	
6	4	1	Yes			
6	4	2	No	$\{\{1, 2\}, \{3, 4, 5, 6\}\}(*)$	No	All partitions (*)
6	4	3	Yes		Yes	
7	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	
8	$\{1, 2, 3\}$	$\{1, \dots, m\}$	Yes		Yes	

Table A.1: Results for the systematic generation of profiles. Note that for $n = 3$, $m \leq 6$. For $n = 4$, $m \leq 5$ and for $n \in [5, 8]$ it is the case that $m \leq 4$. (*): Counterexamples to EJRs being satisfied were found, but exact percentages are unknown.